

09/6/13 486

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein: - protein search, using sw model

Run on: November 7, 2003, 14:47:43 ; Search time 21 Seconds  
(without alignments)  
398.931 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMSDSNLSNLVITDASSL.....GGVNTVPVSNLRQLGRREVM 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 423.0858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*

1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/6C\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/6D\_COMB.pep.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARY-ES

Result No.	Score	Query Match	Length	ID	Description
1	99.1	100.0	198	3	US-09-080-983-15
2	85.6	8.6	593	2	US-09-910-551B-2
3	8.2	732	4	US-09-252-991A-30074	Sequence 2, Appl
4	79.5	8.0	770	1	US-08-445-135-2
5	79.5	8.0	477	2	US-08-770-544-16
6	77.5	7.8	410	4	US-09-252-991A-20306
7	77.5	7.8	410	4	US-09-252-991A-20306
8	77.5	7.8	537	3	US-09-252-991A-32657
9	76.7	7.7	223	3	US-09-080-983-13
10	74.5	7.5	587	4	US-09-107-532A-6405
11	73.5	7.4	706	1	US-08-484-105-16
12	73.5	7.4	706	1	US-08-484-105-16
13	73.5	7.4	503	3	US-09-357-251-36
14	72.5	7.3	641	3	US-09-422-869-26
15	72.5	7.3	553	6	5310678-1
16	72.5	7.3	707	2	US-08-576-165-4
17	71.5	7.2	486	4	US-09-252-991A-31879
18	71.5	7.2	597	4	US-09-252-991A-21752
19	71.5	7.2	1048	4	US-09-171-699-10
20	71.5	7.2	193	4	US-08-671-548C-46
21	71.5	7.2	1040	4	US-09-328-352-7238
22	70.5	7.1	359	1	US-08-137-627-4
23	70.5	7.1	359	2	US-08-865-348-4
24	70.5	7.1	435	2	US-08-531-439B-4
25	70.5	7.1	586	4	US-09-107-532A-6918
26	69.5	7.0	206	4	US-09-107-532A-6952
27	69.5	7.0	414	4	US-09-252-991A-17176

Sequence 24, Appl  
Sequence 24, Appl  
Sequence 6, Appl  
Sequence 2, Appl  
Sequence 13, Appl  
Sequence 4807, Ap  
Sequence 6865, Ap  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 11, Appl  
Sequence 5, Appl  
Sequence 2, Appl  
Sequence 4377, Ap  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-080-983-15  
; Sequence 15, Application: US/09080983  
; Patent No. 6197948  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Hai-Ying  
; APPLICANT: Ling, Kai-Shu  
; APPLICANT: Gonsalves, Dennis  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS  
; TITLE OF INVENTION: AND THEIR USES  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/080,983  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/047,194  
; FILING DATE: 20-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1631  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-080-983-15

Query Match: 100.0%; Score 991; DB 3; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.7e-105;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 XELMSDSNLSNLVITDASSLNGVCKLLSAEVEKMLVQKAGNEGIEVVGGLLYALAAR 60  
DB 1 MEMSDSNLSNLVITDASSLNGVCKLLSAEVEKMLVQKAGNEGIEVVGGLLYALAAR 60  
QY 61 TTSPKQVQADSDVIFSNSFGERNVVVTEGDLKKVLDGCAPLFRFTNKLRITFGRTTEAYV 120  
DB 61 TTSPKQVQADSDVIFSNSFGERNVVVTEGDLKKVLDGCAPLFRFTNKLRITFGRTTEAYV 120  
QY 121 DFCIAYKHLPOLNAAAEIGIPAEDSYLAADFLGTCPKLSELQQRKMPASMYALKTEGG 180  
DB 121 DFCIAYKHLPOLNAAAEIGIPAEDSYLAADFLGTCPKLSELQQRKMPASMYALKTEGG 180  
QY 181 VNTPVSNLRQLGRREV 198  
DB 181 VNTPVSNLRQLGRREV 198

## RESULT 2

US-08-910-551B-2  
Sequence 2, Application US/08910551B  
Patent No. 5910571  
GENERAL INFORMATION:  
APPLICANT: Sylvain Moineau, Barbara  
APPLICANT: J. Holler, Peter A. Vandenbergh,  
APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.  
APPLICANT: Kondo  
TITLE OF INVENTION: DNA Encoding Phage  
TITLE OF INVENTION: Abortive Infection Protein  
TITLE OF INVENTION: From Lactococcus  
TITLE OF INVENTION: Lactis, and Method of Use Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Ian C. McLeod  
STREET: 2190 Commons Parkway  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette 5.25 inch,  
MEDIUM TYPE: 360 Kb storage  
COMPUTER: Acer  
OPERATING SYSTEM: MS-DOS (version: 4)  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,551B  
FILING DATE: August 13, 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/565,907  
FILING DATE: December 1, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: Quest 4.1-158  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
TELEX: No. 591057-e  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 599  
TYPE: Amino Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: No  
HYPOTHETICAL: No  
ANTI-SENSE: No  
FRAGMENT TYPE: N/A  
ORIGINAL SOURCE:  
ORGANISM: Lactococcus lactis  
STRAIN:

INDIVIDUAL ISOLATE: W1  
DEVELOPMENTAL STAGE: N/A  
HAPLOTYPE: N/A  
TISSUE TYPE: N/A  
CELL TYPE: Bacterium  
CELL LINE: N/A  
ORGANELLE: N/A  
IMMEDIATE SOURCE:  
LIBRARY: genomic  
CLONE: SMQ-20  
POSITION IN GENOME: N/A  
FEATURE:  
NAME/KEY: phage abortive infection  
LOCATION: N/A  
IDENTIFICATION METHOD: based upon DNA  
OTHER INFORMATION: phage resistance  
OTHER INFORMATION: enzyme Ab-E  
PUBLICATION INFORMATION: N/A  
US-08-910-551B-2

Query Match 8.8%; Score 85; DB 2; Length 599;  
Best Local Similarity 24.3%; Pred. No. 0.5;  
Matches 36; Conservative 26; Mismatches 46; Indels 40; Gaps 6;  
QY 2 ELMDSNL---SNLVTD-----ASSLNGVDK--KLSA 30  
DB 261 EFNENFLICRENNLIINDNKTKVDNPPFVDSKSDIFSFFENITSTNSCKWKEISN 320  
QY 31 EVELMLVQKAGNEGIEVVGGLLYALAARTTSPKQVQADSDVIFSNSFGERNVVVTEG 89  
DB 321 FIDYCVNEEHLGNKAIGKIFPVI-----TNLKGKVDTKNI-DN-PSKSNVVTNPN 372  
QY 90 DLKKVLDGCAPLFRFTNKLRITFGRTTE 117  
DB 373 VFELDLSLKOSRLTNKFLTFPENINE 400

## RESULT 3

US-09-252-991A-30074  
Sequence 30074, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 30074  
LENGTH: 732  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30074  
Query Match 8.2%; Score 81; DB 4; Length 732;  
Best Local Similarity 22.4%; Pred. No. 1.9;  
Matches 30; Conservative 26; Mismatches 50; Indels 28; Gaps 5;  
QY 17 ASSLNGVDKLLSAEVEKMLVQKAGNEGIEVVGGLLYALAARTTSPKQVQADSDVIFS 76  
DB 44 ANTMGVYREAMAKTIAELEAEK-----GVVL-----TSAKKTFAGGDL--- 87  
QY 77 NSFGERNVVVTEGDLKKVLDGCAPLFRFTNKLRITFGRTTEAY-----VDFCIAYKH 128  
DB 88 ----NELIKVTADAPAFYQIGILEKQJLRRLTLGKPPVVAALNGAALGGGWEICLACHH 143  
QY 129 KLPOLNAAAEIGIP 142

Db 144 RIALDNPVQLGLP 157

## RESULT 4

US-08-445-135-2  
; Sequence 2, Application US/08445135  
; Patent No. 5658789  
; GENERAL INFORMATION:  
; APPLICANT: Quaranta, Vito  
; TITLE OF INVENTION: Promotion of Epithelial Cell Adhesion  
; ADDRESS: Hormia, Marketta  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/445,135  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/151,134  
; FILING DATE: 12-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: DESXOS.002CP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 770 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; IMMEDIATE SOURCE:  
; CLONE: 150 kD  
US-08-445-135-2

Query Match 8.0%; Score 79.5; DB 1; Length 770;

Best Local Similarity 26.6%; Pred. No. 3.1;  
Matches 38; Conservative 22; Mismatches 54; Indels 29; Gaps 7;

Qy 2 ELMSDLSNLVITDA--SSINGVDKLLSAEVEKMLVOKGAPNEGIEVVGGLLVALAA 59  
Db 325 ELLEAKWTRKLCQCEINPALNSLOQTLKTVSVCKLLDAN-----VTAV 369  
Qy 60 RTTSPKVQADSDVIFNSGGERNVV-VTEGDLKKYLDGQCAPJTRFTNKLRT-FGRT--- 114  
Db 370 RNDLRGIQRGIDSWVS---GAKSMVRKANGITSEVLDSAPSRRRIWEGIRATGCTRHE 426  
Qy 115 -FTEAYVDFCIAYK---HKLPL 133  
Db 427 DFNKALIDANNSVKLTKCPDL 449

## RESULT 5

US-08-770-544-16  
; Sequence 16, Application US/08770544

; Patent No. 5907085  
; GENERAL INFORMATION:  
; APPLICANT: Gonsalves, Dennis  
; APPLICANT: Ling, Kai-Shu  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND  
; TITLE OF INVENTION: THEIR USES  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/770,544  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60009008  
; FILING DATE: 21-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael J.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/621  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 477 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-770-544-16

Query Match 8.0%; Score 79; DB 2; Length 477;

Best Local Similarity 21.0%; Pred. No. 1.8;  
Matches 43; Conservative 40; Mismatches 60; Indels 62; Gaps 11;

Qy 12 LVITDASSLNGVDKLLSAEVEKMLVOKG---APNE-GIEVVGGLLYALAAATTSPKVC 67  
Db 302 LPITTEALQIN-----ARLRKVLKSGSGQTPRDXGNMIVAMIQLFVLYSTVKXISVK 353  
Qy 68 ---RADSDVIFNSGGERNVVVTGDLKKYL---DGCAPLRTFTNKLRTFGRTTEAYV 120  
Db 354 DGYRVETEL-----GOKRVLYSYSEVREAILGKYGASP-----TNTVRSFMYFAHTT 403  
Qy 121 DFCIAYKHKLQPLNAAAEGLIP-----AEDSYLAADFLGTCPKLSELQSRKM 168  
Db 454 TLLIEKKIQ-PACTALAKHGVPKRFTPCDFALLDNRYPADVLT-----KANA 451  
Qy 169 FASMYALKTEGGVVNTPVSNLRGLG 193  
Db 452 MACAIAIKS-----ANLRKKG 467

## RESULT 6

US-09-579-259-16  
; Sequence 16, Application US/09579259  
; Patent No. 6558953  
; GENERAL INFORMATION:  
; APPLICANT: Gonsalves, Dennis  
; APPLICANT: Ling, Kai-Shu  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS  
; PROTEINS AND THEIR USES  
; NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
 STREET: Clinton Square, P.O. Box 1051  
 CITY: Rochester  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 14603

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/579,259  
 FILING DATE: 25-May-2000  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 6009008  
 FILING DATE: 21-DEC-1995

ATTORNEY/AGENT INFORMATION:  
 NAME: Goldman, Michael L.  
 REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/621

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (716) 263-1304  
 TELEFAX: (716) 263-1603

INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 477 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-09-579-259-16

Query Match 8.0%; Score 79; DB 4; Length 477;

Best Local Similarity 21.0%; Pred. No. 1.8;

Matches 43; Conservative 40; Mismatches 60; Indels 62; Gaps 11;

QY 12 LVITDASSLVGVKLLSAEVEKMLVQKQ---APNE-GIEVVGILLVALAARTTSPKVV 67

Db 302 LPITEALQIN-----AELRLVLSKSGSOTPRDMGNVIANIQLFVLYSTVKVNSVK 353

QY 68 ---RADSDVFSNGFERNVWVTEGDKKVL---DGCAPLTFRTNKLTFCRTTEAVV 120

Db 354 DGVRYVETEL-----GQKRVY:SYSEVREAILGKYGASP-----TNTVRSFNRYFAHTTI 403

QY 121 DFCIAYKHKLQVNAAEELGIP-----AECSYLAADFLGTCPKLSELOQSRKM 168

Db 404 TLLIEKKIQ-PACTALAKHGVPKRFTFYCFDFALLDNRYYPADVL-----KANA 451

QY 169 FASMYALKTEGGVWVTPVSNLRCLG 193

Db 452 MACATAIKS-----ANLRKRG 467

RESULT 7

US-09-252-991A-20306

Sequence 20306, Application US/09/252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 07196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/374,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/394,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

Query Match 7.8%; Score 77; DB 4; Length 597;

Best Local Similarity 34.0%; Pred. No. 4.1;

Matches 33; Conservative

QY 99 APLTRFTNKLTFRTRTEAYVDFC-IAYYKHLQVNAAEELGIPAEEDSYLAADF-GTCTP 157

Db 461 AALFRGTTELLHPSGSTILEADDILCVIGHEDLP---ALGKLFSQA?DRGLGARFFGDFV 517

QY 158 KLSELOQSRKMFASMYALKTEGGVWVTPVSNLRCLGR 194

Db 519 LEGDAQLS--AVASLYGLKLDG-----IDGEOALGR 546

RESULT 9

US-09-080-983-13

Sequence 13, Application US/09080983

Patent No. 6197948

GENERAL INFORMATION:

APPLICANT: Zhu, Hai-Ying

APPLICANT: Ling, Kai-Shu

APPLICANT: Gonsalves, Dennis

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS

NUMBER OF SEQUENCES: 23

SEQ ID NO 20306

LENGTH: 410

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20306

Query Match 7.8%; Score 77.5; DB 4; Length 410;

Best Local Similarity 25.0%; Pred. No. 2.1;

Matches 60; Conservative 31; Mismatches 90; Indels 59; Gaps 13;

QY 7 SNLSNLVITDASSLVGVKLLSAEVEKMLVQKQ-----GAPNEGIEVV 49

Db 78 SCMDGLAFURHASLSGKVVHSLSSSEVDP-LRQAT?SMIECLGLNFLGDLGKP-FS?ER- 136

QY 50 FGLLYVALAARTTSPK-VQRAD-----SDVIFSNSPGE-----RNVVTEGDL--KKVLD 96

Db 137 TALLTRYNNARQDLPRQ?EVAELPSVADVVRGJDNGEFAVYQPKVALDGGGLIGAEVLA 196

QY 97 -----GCAPLTFRTNKLT-----FGRTTTEAYVDFC-IAYYKHLQVNAAEELG-P 142

Db 197 RWRHPHLGVPPSHFLYVMETYNLVKLFWQLFSQG-----CATRRKLAQLGQPINLAFN 251

QY 143 AEDSYLAADF--GTCPKLSE--LQSRKMFASMYALKTEGGVWVTPVSNLRCLGRREV 198

Db 252 VHPQLGSRALAEINISALLTEFHLPPSSVMFEI-----TETGLISAPASSLENLVRLR-M 306

RESULT 8

US-09-252-991A-32657

Sequence 32657, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 32657

LENGTH: 597

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-32657

Query Match 7.8%; Score 77; DB 4; Length 597;

Best Local Similarity 34.0%; Pred. No. 4.1;

Matches 33; Conservative

QY 99 APLTRFTNKLTFRTRTEAYVDFC-IAYYKHLQVNAAEELGIPAEEDSYLAADF-GTCTP 157

Db 461 AALFRGTTELLHPSGSTILEADDILCVIGHEDLP---ALGKLFSQA?DRGLGARFFGDFV 517

QY 158 KLSELOQSRKMFASMYALKTEGGVWVTPVSNLRCLGR 194

Db 519 LEGDAQLS--AVASLYGLKLDG-----IDGEOALGR 546

RESULT 9

US-09-080-983-13

Sequence 13, Application US/09080983

Patent No. 6197948

GENERAL INFORMATION:

APPLICANT: Zhu, Hai-Ying

APPLICANT: Ling, Kai-Shu

APPLICANT: Gonsalves, Dennis

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS

NUMBER OF SEQUENCES: 23



;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
;/ STREET: Clinton Square, P.O. Box 1051  
;/ CITY: Rochester  
;/ STATE: New York  
;/ COUNTRY: U.S.A.  
;/ ZIP: 14603  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: Patent'n Release #1.0, Version #1.30  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/09/680,983  
;/ FILING DATE:  
;/ CLASSIFICATION:  
;/ PRIOR APPLICATION NUMBER: US 60/047,194  
;/ FILING DATE: 20-MAY-1997  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Goldman, Michael L.  
;/ REGISTRATION NUMBER: 30,727  
;/ REFERENCE/DOCKET NUMBER: 19603/163:  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (716) 263-1304  
;/ TELEFAX: (716) 263-1600  
;/ INFORMATION FOR SEQ ID NO: 13:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 223 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS:  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: protein  
;/ US-09-080-983-13

Query Match 7.7%; Score 76; DB 3; Length 223;  
Best Local Similarity 23.3%; Pred. No. 1.3;  
Matches 48; Conservative 36; Mismatches 88; Indels 34; Gaps 11;  
QY 4 MDSNLSNLVITDASSLNGVD-KLLSAEVEKMLVQKAPKEGIEVVGGLLYAL--AAR 60  
DB 36 LKNENYSSV---DSSRLSDSEVKEVLEKSEKSFSELASTDE--HFVYHIFFLIRCAKI 90  
QY 61 TTSKPVQRAEDSDVIFSNFSGERNVVVTEGDLKKVLDG-----CAPLT---RFTNKLRTFG 112  
DB 91 STSEKVKVGS-----H-YVVDGKTYVLDWVFNMMKSLTKYKRVNGLRAFEC 139  
QY 133 RTFTTEAYVDFCIAYKHKLPOLNMAAELGIPAEDSYLADEL--GTCPKLSLQSQSRKMFAS 172  
DB 140 CACEDLYLTVPAPINSERF-KTKAVOMKGLPVGKEVYLGADFLSGTSLKMSCHDRAVSIVAA 198  
QY 172 MYALKTE---GGVNTVPVSNLRQLQR 194  
DB 199 KNAVDRSAFTGG--ERKIVSLYDLGR 222  
RESULT 10  
US-09-107-532A-6405  
; Sequence 6405, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:

;/ MEDIUM TYPE: CD-ROM, ISO9660  
;/ COMPUTER: PC  
;/ OPERATING SYSTEX: <Unknown>  
;/ SOFTWARE: ASCII  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/09/107,532A  
;/ FILING DATE: 30-Jun-1998  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: 60/085,598  
;/ FILING DATE: 14 May 1998  
;/ APPLICATION NUMBER: 60/051571  
;/ FILING DATE: July 2, 1997  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Ariniello, Pamela Deneke  
;/ REGISTRATION NUMBER: 40,489  
;/ REFERENCE/DOCKET NUMBER: GTC-012  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (781)893-5007  
;/ TELEFAX: (781)893-8277  
;/ INFORMATION FOR SEQ ID NO: 6405:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 587 amino acids  
;/ TYPE: amino acid  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: protein  
;/ HYPOTHETICAL: YES  
;/ ORIGINAL SOURCE:  
;/ ORGANISM: Enterococcus faecium  
;/ FEATURE:  
;/ NAME/KEY: misc feature  
;/ LOCATION: (B) LOCATION 1...587  
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 6405:  
US-09-107-532A-6405  
Query Match 7.5%; Score 74.5; DB 4; Length 587;  
Best Local Similarity 22.6%; Pred. No. 7.8;  
Matches 43; Conservative 27; Mismatches 65; Indels 55; Gaps 10;  
QY 1 MELSDSNLSNLVITDASS---LNGVDEKXLLSAEVEKMLVCKGAPNEG--EVVVF-----50  
DB 224 MEESGKVLNHLXNAENKTETFTGIGNKQKQKQEVAKLIVE-----NFMKQVFDGFFHA 278  
QY 51 ---GLLYALAAATTSKPVQRAEDSDVIFSNFSGE---RNWVTEGDLKKVLDGCAPLTR 103  
DB 279 DHPGNLLFHLVLTKEEQTCASR-KTETVHEKEFGSPAFRASTSAE-----DVPAP---327  
QY 104 FTNKLRTFGRTFTAYVDF-----CIAYKHKLPOLNMAAELGIPAEDSY-LAADELSTC 156  
DB 328 -----YTINYIDFGWGHLSAGLRQKLTQ-----AVLALYTKDAYRIEKAVALRLC 372  
QY 157 PKLSLQSQSR 166  
DB 373 QQEGSFDES 382  
RESULT 11  
US-08-484-105-16  
; Sequence 16, Application US/08484105  
; Patent No. 5589341  
; GENERAL INFORMATION:  
; APPLICANT: STILLMAN, Bruce  
; APPLICANT: BELL, Stephen P  
; APPLICANT: KOBAYASHI, Ryuji  
; APPLICANT: RINE, Jasper  
; APPLICANT: FOSS, Margit  
; APPLICANT: MCNALLY, Francis J  
; APPLICANT: LAURENSEN, Patricia  
; APPLICANT: HERSKOWITZ, Ira  
; APPLICANT: LI, Joachim J  
; APPLICANT: GAVIN, Kimberly  
; TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DCS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,105  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard Aron  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59032/DJE/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 706 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-105-16

Query Match 7.4%; Score 73.5; DB 1; Length 706;  
Best Local Similarity 19.3%; Pred. No. 13;  
Matches 43; Conservative 33; Mismatches 68; Indels 79; Gaps 9;

Qy 10 SNLVITDASSINGVDKLLSAEVEKMLVQKGAPE-----GIBVFGLLLYALAARTTSP 64  
Db 479 SRLIVAVANTMDLPRLNSRISRLGLSRVFPFYTHQTLEII-----IAARLEAV 531  
Qy 65 KVRADSDVIFSNS--FGERNVVVTEDLKKVLDGC----- 98  
Db 532 R-----DDVFSDDA:RFAARKVAASGDARRALD:ICRRASELAENKNGKVTPLIHQAIS 587  
Qy 99 -----APLTR-----FTNKLRTFGRTTEAYVDFCIAYKHKLPOLNAAEL 139  
Db 588 EM:ASPLQKVLRLNLSFMQKVFCLCAIVNMRMSG--FAESVYVEVLEAEKRLRVMTTPD- 644  
Qy 140 GIPAEYSYLAADFLGTCPKLSLQSRKMFASVYALKTEGGVV 192  
Db 645 ---AE-----AKFGELILRRPFGYVLSLSSENGVL 672

RESULT 12  
US-08-484-106-16  
Sequence 16, Application US/08484106  
Patent No. 5614618  
GENERAL INFORMATION:  
APPLICANT: STILLMAN, Bruce  
APPLICANT: BELL, Stephen P  
APPLICANT: KOBAYASHI, Ryuji  
APPLICANT: RINE, Casper  
APPLICANT: FOSS, Margit  
APPLICANT: McNALLY, Francis J  
APPLICANT: LAURENSEN, Patricia  
APPLICANT: HERSKOWITZ, Ira  
APPLICANT: LI, Joachim J  
APPLICANT: GAVIN, Kimberly  
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DCS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,106  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman Ph.D., Richard Aron  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 706 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-106-16

Query Match 7.4%; Score 73.5; DB 1; Length 706;  
Best Local Similarity 19.3%; Pred. No. 13;  
Matches 43; Conservative 33; Mismatches 68; Indels 79; Gaps 9;

Qy 10 SNLVITDASSINGVDKLLSAEVEKMLVQKGAPE-----GIBVFGLLLYALAARTTSP 64  
Db 479 SRLIVAVANTMDLPRLNSRISRLGLSRVFPFYTHQTLEII-----IAARLEAV 531  
Qy 65 KVRADSDVIFSNS--FGERNVVVTEDLKKVLDGC----- 98  
Db 532 R-----DDVFSDDA:RFAARKVAASGDARRALD:ICRRASELAENKNGKVTPLIHQAIS 587  
Qy 99 -----APLTR-----FTNKLRTFGRTTEAYVDFCIAYKHKLPOLNAAEL 139  
Db 588 EM:ASPLQKVLRLNLSFMQKVFCLCAIVNMRMSG--FAESVYVEVLEAEKRLRVMTTPD- 644  
Qy 140 GIPAEYSYLAADFLGTCPKLSLQSRKMFASVYALKTEGGVV 192  
Db 645 ---AE-----AKFGELILRRPFGYVLSLSSENGVL 672

RESULT 13  
US-09-357-251-36  
Sequence 36, Application US/09357251  
Patent No. 6271441  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
APPLICANT: Pamodu, Zayo O.  
APPLICANT: Orozco, Buddy  
APPLICANT: Schwaber, James S.  
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase  
FILE REFERENCE: BB-1193  
CURRENT APPLICATION NUMBER: US/09/357,251  
CURRENT FILING DATE: 1999-07-20  
EARLIER APPLICATION NUMBER: 60/093,530  
EARLIER FILING DATE: July 21, 1998  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 36  
LENGTH: 503  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-357-251-36

Query Match 7.4%; Score 73; DB 3; Length 503;  
Best Local Similarity 24.6%; Pred. No. 9.2;  
Matches 48; Conservative 33; Mismatches 74; Indels 40; Gaps 12;  
  
QY 4 MDSNLSNLV-ITDASSLNGVDKLLSAEVEKMLVQKG-----APNEGIEVVFSLLYAL 57  
DB 134 LTDTQTILAIQIKNSHSDSIDAKILNDJKRKLROGKITDFSVTGPE--FSTDLTKL 191  
QY 58 AARTSPKVQ-RADSDVIFS-NSFGERNVVVTEGLKVLGCGAPLRTFNKLT--FGR 113  
DB 192 ETCJTSQVSTNAYKDKLPYFNFSQGVQISSGALH-----PLNKVREERQIFFSM 244  
QY 114 TTFE-----AYVD-----FCIAYKKLQOLNAAELGIPAEDSVLAADFLCTCPKLSLQ 164  
DB 245 GFTEPMNQVETGFWFDALY---VPOQHPARDL-----QDTFYKD-----ELTAEJFD 292  
QY 165 SRKYFASMWALKTEG 179  
DB 293 DKTYMONIKAVHEQG 307

RESULT 14  
US-09-422-969-26  
; Sequence 26, Application US/09422869  
; Patent No. 6235481  
; GENERAL INFORMATION:  
; APPLICANT: POLONSKI, KENNETH S.  
; APPLICANT: HORIKAWA, YUKIO  
; APPLICANT: ODA, NACHISA  
; APPLICANT: COX, NANCY J.  
; APPLICANT: SREEMAN, SEAMUS  
; APPLICANT: ZHOC, YUN-PING  
; APPLICANT: OTANI, KENICHI  
; APPLICANT: HANIS, CRAIG L.  
; APPLICANT: BELL, GRAEME I.  
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES  
; FILE REFERENCE: ARCD:307  
; CURRENT APPLICATION NUMBER: US/09/422,969  
; CURRENT FILING DATE: 1999-10-21  
; EARLIER APPLICATION NUMBER: 60/134,175  
; EARLIER FILING DATE: 1999-05-13  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 641  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-422-869-26

Query Match 7.3%; Score 72.5; DB 3; Length 641;  
Best Local Similarity 25.4%; Pred. No. 15;  
Matches 29; Conservative 19; Mismatches 43; Indels 23; Gaps 6;  
  
QY 79 FGERNVVTEGLKKVLGCGAPLRTFNKLTFTGRTFE-AYVDFCIAYKHKLPLNAAA 137  
DB 128 FGEWTEVVID-DLLPTINGDLVFS-FSTSMNPNALLERKAYAK-LGQYE-----A 176  
QY 138 ELGIPAEDSVLAADFLCTCPKLSLQSR-----KMFASNYALKTEGGVV 182  
DB 177 LDGLTITD--IIMDPTGTLABEIDMQGRYTDLVEEKYKLFGLYKRTFKGGLI 228

RESULT 15  
5310678-1  
; Patent No. 5310678  
; APPLICANT: Bingham, Richard W.; Chambers, Philip; Emerson, Peter  
; T. Millar, Neil S.  
; TITLE OF INVENTION: NEWCASTLE DISEASE VIRUS GENE CLONES  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/438,945  
; FILING DATE: 17-NOV-1989

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 885,765  
; FILING DATE: 15-JUL-1986  
; SEQ ID NO: 1  
; LENGTH: 553  
5310678-1  
  
Query Match 7.3%; Score 72; DB 6; Length 553;  
Best Local Similarity 24.1%; Pred. No. 14;  
Matches 49; Conservative 28; Mismatches 94; Indels 32; Gaps 8;  
  
QY 3 LMSDSNLSNLVITDASSLN-GVCKKLLSAEVEKMLVCKGAPNEGIEVVF---GLLYAL 57  
DB 289 LPSVGNLNNMRATYLETLSVSTTRGFASALVPKVTQVGSVIELDTSYCIETOLDLY-- 346  
QY 58 AARTSPKVQRAADSEVIFPSNFGERNVVV---TEGDLK-----KVLGCGAPLT-RFT 105  
DB 347 CTRIVTPFM---SPQIYSCLSGNTSACMYSKTEGALTTPYMTIKGSVIANCKMTTCRCV 402  
QY 106 NKLRTEGRTFTTEAYVDPCIAKHKLPOLNAAELGIPAEDSVLAADFLCTCPKLSLQCS 165  
DB 403 NPPGIISQNYGEAV-----SLDPKQSCNVLSDGITLRLSGEPDATYCKNISIQCS 453  
QY 166 RKMFASMYALKTEGGVNTVPVSN 188  
DB 454 QVIITGNLIDISTELGNVWNSISN 476

Search completed: November 7, 2003, 14:50:55  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2003, 14:47:43 ; Search time 21 seconds  
(without alignments)  
398.93; Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 MELASDSNLSNLVITDASSL.....GGVNVTPVSNRLQLGRREV 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA\*

1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	198	3	US-09-080-983-15
2	85	8.6	599	2	US-08-910-551B-2
3	81	8.2	732	4	US-09-252-991A-30074
4	79.5	8.0	770	1	US-08-445-135-2
5	79	8.0	477	2	US-08-770-544-16
6	79	8.0	477	4	US-09-579-259-16
7	77.5	7.8	410	4	US-09-252-991A-20306
8	77	7.8	597	4	US-09-252-991A-32657
9	76	7.7	223	3	US-09-080-983-13
10	74.5	7.5	587	4	US-09-107-532A-6405
11	73.5	7.4	706	1	US-08-484-105-16
12	73.5	7.4	706	1	US-08-484-106-16
13	73	7.4	503	3	US-09-357-251-36
14	72.5	7.3	541	3	US-09-422-869-26
15	72	7.3	553	6	5310678-1
16	72	7.3	707	2	US-08-576-185-4
17	71.5	7.2	486	4	US-09-252-991A-31879
18	71.5	7.2	597	4	US-09-252-991A-21752
19	71.5	7.2	1048	4	US-09-171-699-10
20	71	7.2	193	4	US-08-671-548C-46
21	71	7.2	1040	4	US-09-328-352-7238
22	70.5	7.1	359	1	US-08-137-627-4
23	70.5	7.1	359	2	US-08-865-348-4
24	70.5	7.1	435	2	US-08-531-439B-4
25	70.5	7.1	586	4	US-09-107-532A-6918
26	69.5	7.0	206	4	US-09-107-532A-6952
27	69.5	7.0	414	4	US-09-252-991A-17176

28 59.5 7.0 1713 3 US-08-600-982-24 Sequence 24, Appl  
29 69.5 7.0 1733 5 PCT-US94-10261A-24 Sequence 24, Appl  
30 69.5 7.0 2042 4 US-09-077-098A-6 Sequence 6, Appl  
31 69.5 7.0 2285 4 US-09-308-375-2 Sequence 2, Appl  
32 69 7.0 363 3 US-08-984-618-13 Sequence 13, Appl  
33 68.5 6.9 382 4 US-09-107-532A-4807 Sequence 4807, Ap  
34 68.5 6.9 398 4 US-09-328-352-6865 Sequence 6865, Ap  
35 68 6.9 693 2 US-08-463-620-11 Sequence 11, Appl  
36 68 6.9 693 2 US-08-463-620-11 Sequence 11, Appl  
37 68 6.9 693 2 US-08-463-620-11 Sequence 11, Appl  
38 68 6.9 693 2 US-08-463-620-11 Sequence 11, Appl  
39 68 6.9 693 2 US-08-463-620-11 Sequence 11, Appl  
40 68 6.9 911 1 US-08-596-585-2 Sequence 2, Appl  
41 67.5 6.8 424 4 US-09-328-352-4377 Sequence 4377, Ap  
42 67.5 6.8 875 1 US-08-460-547A-10 Sequence 10, Appl  
43 67.5 6.8 875 1 US-08-460-547A-10 Sequence 10, Appl  
44 67.5 6.8 875 1 US-08-463-949A-10 Sequence 10, Appl  
45 67.5 6.8 875 3 US-08-464-410A-10 Sequence 10, Appl

#### ALIGNMENTS

RESULT: 1  
US-09-080-983-15  
; Sequence 15, Application US/09080983  
; Patent No. 6,979,488  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Hai-Ying  
; APPLICANT: Ling, Kai-Shu  
; APPLICANT: Gonsalves, Dennis  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS  
; TITLE OF INVENTION: AND THEIR USES  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentir Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/080,983  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/047,194  
; FILING DATE: 20-MAY-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/1631  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-080-983-15

Query Match 100.0%; Score 99.1; DB 3; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.7e-105;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2003, 14:49:54 ; Search time 23 Seconds

(without alignments)  
1172.633 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 MELXSDNSLNVLVTDASS.....GGVNTPVSNLRLGRREV 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	286	28.9	204	12	US-10-138-842A-36
2	120.5	12.2	223	12	US-10-138-842A-37
3	79	8.0	451	9	US-09-815-242-10721
4	79	8.0	477	12	US-10-138-842A-16
5	79	8.0	477	12	US-10-032-585-7808
6	77	7.8	696	12	US-10-156-761-11789
7	76.5	7.7	356	15	US-10-156-761-11789
8	75	7.6	2626	10	US-09-799-508B-2
9	74.5	7.5	292	15	US-10-156-761-10849
10	74	7.5	368	15	US-10-205-823-377
11	74	7.5	399	9	US-09-925-301-1385
12	73.5	7.4	451	15	US-10-156-761-12518
13	73.5	7.4	833	9	US-09-844-281-1
14	73	7.4	485	15	US-10-156-761-13439
15	73	7.4	1987	12	US-10-032-585-7518

Sequence 2, Appli.  
Sequence 26, Appli.  
Sequence 18, Appli.  
Sequence 7333, Ap  
Sequence 10, Appli  
Sequence 3, Appli  
Sequence 5013, Ap  
Sequence 10878, A  
Sequence 16, Appli  
Sequence 14, Appli  
Sequence 12853, A  
Sequence 5570, Ap  
Sequence 12280, A  
Sequence 6806, Ap  
Sequence 3, Appli  
Sequence 113, App  
Sequence 6, Appli  
Sequence 2, Appli  
Sequence 7, Appli  
Sequence 394, App  
Sequence 666, App  
Sequence 8805, Ap  
Sequence 7360, Ap  
Sequence 42, Appli  
Sequence 1013, Ap  
Sequence 1934, Ap  
Sequence 1, Appli  
Sequence 1934, Ap  
Sequence 1934, Ap  
Sequence 5, Appli

16 72.5 7.3 183 9 US-09-976-451-2  
17 72.5 7.3 641 10 US-09-768-877-26  
18 72 7.3 707 12 US-10-087-402-18  
19 72 7.3 1119 12 US-10-032-585-7333  
20 71.5 7.2 1048 15 US-10-223-538-10  
21 70.5 7.1 184 9 US-09-976-451-3  
22 70.5 7.1 584 9 US-09-815-242-5013  
23 70.5 7.1 589 9 US-09-815-242-10878  
24 70 7.1 365 15 US-10-125-692-16  
25 70 7.1 452 11 US-09-770-509-14  
26 69.5 7.0 522 15 US-10-156-761-12853  
27 69.5 7.0 530 9 US-09-815-242-5570  
28 69.5 7.0 530 9 US-09-815-242-12280  
29 69.5 7.0 720 10 US-09-738-626-6806  
30 69.5 7.0 833 10 US-09-891-216-3  
31 69.5 7.0 1713 15 US-10-171-311-113  
32 69.5 7.0 2042 15 US-10-192-584-6  
33 69.5 7.0 2285 10 US-09-932-183A-2  
34 69 7.0 363 11 US-09-829-275-7  
35 69 7.0 803 10 US-09-801-268-394  
36 69 7.0 1610 10 US-09-738-626-6666  
37 68.5 6.9 285 15 US-10-156-761-8805  
38 68.5 6.9 526 12 US-10-032-585-7360  
39 68 6.9 360 10 US-09-873-980-42  
40 68 6.9 435 12 US-10-238-075-1013  
41 68 6.9 464 10 US-09-902-941-1934  
42 68 6.9 464 12 US-09-747-524-1  
43 68 6.9 464 12 US-10-113-872-1934  
44 68 6.9 464 15 US-10-017-754-1934  
45 68 6.9 599 15 US-10-138-075-5

# ALIGNMENTS

RESULT 1  
US-10-138-842A-36  
; Sequence 36, Application US:/0138842A  
; Publication No. US20030148390A1  
; GENERAL INFORMATION:  
; APPLICANT: GONSALVES, DENNIS  
; APPLICANT: LING, KAI-SHU  
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND  
; TITLE OF INVENTION: THEIR USES  
; FILE REFERENCE: 07678/025006  
; CURRENT APPLICATION NUMBER: US/10/138,842A  
; CURRENT FILING DATE: 2002-05-03  
; PRIOR APPLICATION NUMBER: US 09/579,259  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 09/224,898  
; PRIOR FILING DATE: 1998-12-31  
; PRIOR APPLICATION NUMBER: US 08/770,544  
; PRIOR FILING DATE: 1996-12-20  
; PRIOR APPLICATION NUMBER: US 60/009,008  
; PRIOR FILING DATE: 1995-12-21  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 204  
; TYPE: PRT  
; ORGANISM: Beet Yellow Virus  
US-10-138-842A-36

Query Match 28.9%; Score 286; DB 12; Length 204;  
Best Local Similarity 37.6%; Pred No. 18e-24;  
Matches 71; Conservative 29; Mismatches 87; Indels 2; Gaps 2;  
Qy 7 SKLSNLVITDASSLNGVDKLLLSAEVEKMLVQKAPNEGIEVVFGLLYALAAATTSPKV 66  
Db 11 ATFNVSADQCLHGEDCDLKRNFECCLKLGVPEDNLGALGLCYSCATIGTSNKV 70  
Qy 67 QRADSDVIFSNF-GERNVVVTEGLKKVLDGCAPLTFRTNKLRTFFSTFTAEVDFCIA 125

```

Dd 71 NYQPTSTFIKASFGGKELYLTHGELNSFLGSKLLEKPKNLRCFCRTFKQDY-SLAKK 130
Qy 126 YHKLPQLNAAELGCPAEDSVLAADFLGTCPKLSELOQRKMFASMYALKTEGGVWTP 185
Dd 131 YRGKLPPIAARHGLPAEDHVLAAAFISTSTELTDLQOSRLLARENATHTEFS-SSSP 189
Qy 186 VSNLRQLGR 194
Dd 190 VTSKQLGR 198

RESULT 2
US-10-138-842A-37
; Sequence 37, Application US/10-138842A
; Publication No. US20030148390A1
; GENERAL INFORMATION:
; APPLICANT: GONSALES, DENNIS
; APPLICANT: LING, KAI-SHU
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
; FILE REFERENCE: 07678/025006
; CURRENT APPLICATION NUMBER: US/10-138,842A
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 09/579,259
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/224,898
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 08/770,544
; PRIOR FILING DATE: 1996-12-20
; PRIOR APPLICATION NUMBER: US 60/009,008
; PRIOR FILING DATE: 1995-12-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Citrus Tristeza Virus
US-10-138-842A-37

Query Match 12.2%; Score 120.5; Db 12; Length 223;
Best Local Similarity 25.8%; Pred. No. 2e-05;
Matches 54; Conservative 31; Mismatches 69; Indels 55; Gaps 10;

Qy 12 LVITDASSINGVDKLLSAEVEKMLVQKG-----APNEG:EVVFGLLLYALAARTS---- 63
Dd 40 ITMNDVRLSTQQAALNRDL--FLAKGKYNLPDKDKDFHAWMLYRLAVKSSLSQSD 97
Qy 64 -----PKVQRAD---SDVIP-SNSFGERNVVTGDLKKVLDGCAPLTRFTNK 107
Dd 98 DDTTGITYTREGVEVDLSOKLWTDIVNSKGIGNR-----TNA 135
Qy 108 LRTFGRTTEAVYDFECAYVGHKLPQLNAAAEGLCPAEDSVLAADFLGTCPKLSELOQRK 167
Dd 136 LRVGRTNDALYLAFCLQNRNLSYSGRPLDAGIPAGYHYLCADFL-TGAGLTDLEAVY 193
Qy 168 MFASMYALKTEGG-VVNTTPVSNLRQLGR 194
Dd 194 IOAKEQLLKKRGADEVV---VTNVRQLGK 215

RESULT 3
US-09-815-242-10721
; Sequence 10721, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
```

```

; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 65/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 65/257,931
; PRIOR FILING DATE: 2003-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10721
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10721

Query Match 8.0%; Score 79; Db 9; Length 451;
Best Local Similarity 24.2%; Pred. No. 3.2;
Matches 53; Conservative 27; Mismatches 73; Indels 66; Gaps 12;

Qy 1 MELMSDSNLS---NLVITDASSL-----NGVDKLLSAEVEKMLVQKGAFNEG:EVVFG 52
Dd 177 LDYAESNAADVDMNIVMTSEGEFVSIQGTGEATFSCDE:NAMLFY---GKTGLEELIAY 233
Qy 53 ---LLYALAARTTSPKVRADSD---V-----FNSNFGERNVVTGDLKKVL 95
Dd 234 QKEALYALA---SEEVPSQDSSEKVIATRNPGKAKEFSS:FGEKGYTV-----KTL 284
Qy 96 DGCAPLTRFTN--KLRTFGRTTEAVYDFECIAYVGHKLPQLNAAAEGLCPAEDSVLAADFL 153
Dd 285 D-----YPNLPDVEETGRTFEB-----NARLKAETIAEILQKPLVADDSGLIVDAL 330
Qy 154 GTCPKLSELOQRKMFASMYALKTEGGVWTPVSNLRQL 192
Dd 331 GGMP-----GIYSARFAGEPTNDASNNAKLL 356

RESULT 4
US-10-138-842A-16
; Sequence 16, Application US/10138842A
; Publication No. US20030148390A1
; GENERAL INFORMATION:
; APPLICANT: GONSALES, DENNIS
; APPLICANT: LING, KAI-SHU
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
; FILE REFERENCE: 07678/025006
; CURRENT APPLICATION NUMBER: US/10-138,842A
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 09/579,259
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/224,898
; PRIOR FILING DATE: 1998-12-31
; PRIOR APPLICATION NUMBER: US 08/770,544
; PRIOR FILING DATE: 1996-12-20
; PRIOR APPLICATION NUMBER: US 60/009,008
; PRIOR FILING DATE: 1995-12-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 477
; TYPE: PRT
```

; APPLICANT: HORIKAWA, HIROSHI



```
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204C89
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11789
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11789

Query Match      7.6%; Score 76.5; DB 15; Length 356;
Best Local Similarity 30.1%; Pred. No. 4.4;
Matches 34; Conservative 12; Mismatches 38; Indels 29; Gaps 6;

QY 57 LAARTSPKVQRADSDVIFSNFSGERNVVTEGDLKKVLDGCAPLRTFTNKLRTFGRTF 115
Db 239 LAARHQAVTTGDPR-LYENG-PEQSQVAIRGAEKADACAAIGRDVKELDKVLTGPF 297

QY 116 T-----EAYVDFCIAYKHKLPOLNAAAEELGI-----PAEDSYLAAD 151
Db 298 TPDGRPLESLDAFVDF-AGRH-----MELGITDIAHKWPIPDSDFAAD 340

RESULT 8
US-09-759-508B-2
; Sequence 2, Application US/09759508B
; Publication No. US20020182599A1
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease
; FILE REFERENCE: 00786/381002
; CURRENT APPLICATION NUMBER: US/09/759,508B
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,787
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 26926
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-508B-2

Query Match      7.6%; Score 75; DB 10; Length 26926;
Best Local Similarity 22.3%; Pred. No. 3.9e+03;
Matches 39; Conservative 27; Mismatches 67; Indels 42; Gaps 8;

QY 54 JYLAARTTS-----PKVQRADSDVI-----FNSNFGERNVVTEGDLKKVLDGCAPL 101
Db 21531 VYTVKAKNSGSAKAEIKVKVQDTPGKWGPFRFTNITGEKMTJWDAFLN---DGCAPL 21587

QY 102 TRTNKLRFTGRTFTAYVDFCI-----YKHKLPLNAAAEELGIPAECSYL 148
Db 21588 THVIIKRETRSLJAWALIEDKCEAQSYATKILNGNEYQFRVSANVKFG-VGRPLSDSFV 21646

QY 149 AADFLGTCPLSELQSRKQKFAWMAL-----KTEGGVWNTPVSNLRQ--LGRRE 196
Db 21647 VAGIQYTVDPAPGIPPSNNTGNSITLTWARPESDGG-----SEIQQVILERR 21695

RESULT 9
US-10-156-761-10849
; Sequence 10849, Application US/10-156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
```

```
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10849
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10849

Query Match      7.5%; Score 74.5; DB 15; Length 292;
Best Local Similarity 28.3%; Pred. No. 5.5;
Matches 39; Conservative 12; Mismatches 48; Indels 33; Gaps 8;

QY 31 EVENMLVQKAPNEGIEVVFGLLYLALAAATTSFKVQRADSDVIFSNFSGERNVVTEGC 90
Db 26 DMHTYLLGGWMDERGAEAVYGFLEAAGQAGR---QRAAGDAAVVGC-----LLVDEGC 77

QY 91 -----LKKVLDDGCAPL-----TRTNKLRFTGRT-----FTBAYVDFCI 124
Db 78 GAGOFARYEAVLRKVAD-CAPVPLLPJSERF--DVGALGRVDALLVCGSLTFAYQDALA 134

QY 125 AYKHKLPLNAAAEELGIP 142
Db 135 EVLGRPLRV--LAERGIP 150

RESULT 10
US-10-205-823-377
; Sequence 377, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 377
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-205-823-377

Query Match
Best Local Similarity 7.5%; Score 74; DB 15; Length 369;
Matches 50; Conservative 36; Mismatches 70; Indels 66; Gaps 12;

QY 24 DKLLSAEVEKMLVQKAGNEGIEV---FGLLLYALAARTTSPKVRADSDVIFSNSF 79
Db 130 DGRLEVDIDEVVDSPQN:KILHSKQFGN:L-----ILSGDVNLAESDLAYTRAI 183

QY 80 -----GERNVVTEG-----LK-----KVLGGCAPLRTFT--NKL 108
Db 184 MSGGKEDYTGKVLILGGGGGGLCEIVKLKPKMVTWVIDQWIDGCKKXKTKCGDVL 243

QY 109 RTFGRTTEAYVDFCI-----AYKHKLPLQNAAAELGIPAEDSYLAADFLGTCP 157
Db 244 DNLKGDCYQVLIEDCI:PVLKRYAKEGREFDYVINDLTAVPIS:SPEDS--TWEFLRLIL 301

QY 158 KLS--ELQQRKMFASMYALKTEGGVNV-TPVSNL--RQLGR 194
Db 302 DLSMKVLKQDGKYP-----TQNCVNLTEALSLEYEQLR 336

RESULT 11
US-09-925-301-1385
; Sequence 1385, Application: US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05982
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/24,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent In Ver. 2.3
; SEQ ID NO 1385
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-301-1385

Query Match
Best Local Similarity 7.5%; Score 74; DB 9; Length 359;
Matches 50; Conservative 36; Mismatches 70; Indels 66; Gaps 12;

QY 24 DKLLSAEVEKMLVQKAGNEGIEV---FGLLLYALAARTTSPKVRADSDVIFSNSF 79
Db 161 DGRLEVDIDEVVDSPQN:KILHSKQFGN:L-----ILSGDVNLAESDLAYTRAI 214

QY 80 -----GERNVVTEG-----LK-----KVLGGCAPLRTFT--NKL 108
Db 215 MSGGKEDYTGKVLILGGGGGGLCEIVKLKPKMVTWVIDQWIDGCKKXKTKCGDVL 274

QY 109 RTFGRTTEAYVDFCI-----AYKHKLPLQNAAAELGIPAEDSYLAADFLGTCP 157
Db 275 DNLKGDCYQVLIEDCI:PVLKRYAKEGREFDYVINDLTAVPIS:SPEDS--TWEFLRLIL 332

QY 158 KLS--ELQQRKMFASMYALKTEGGVNV-TPVSNL--RQLGR 194
Db 333 DLSMKVLKQDGKYP-----TQNCVNLTEALSLEYEQLR 367

RESULT 12
US-10-156-761-12518
; Sequence 12518, Application US/10156761
; Publication No. US20030113018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATCSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HIROSHI, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12518
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-12518

Query Match
Best Local Similarity 7.4%; Score 73.5; DB 15; Length 451;
Matches 42; Conservative 33; Mismatches 60; Indels 61; Gaps 10;

QY 10 SNLVITDASSLNGVDKLLSA-EVERKMLVQKAGNEGIEVFGLLLYALAARTTSPKVR 68
Db 138 AKLYIND-YNVEGVNAKSTALYNLVKSLKRGVPI:DGVLQAHVLVGQVPS-TLQONIOR 255

QY 69 -ADSDVIFSNGSERNVVTE-----GSLKKVLGSCAPLRTFTNKL 109
Db 256 FADLGV-----DVAITELDIRMQLPATDKLTQOADYKAVLDCVAVTRCVG-VT 305

QY 110 TFGRTTEAYVDFCIAYKHKLPLQNAAAELG:PAEDSYLAADF:GTGFKLSELOQSRMF 169
Db 306 VMGFTDSDSWI-----PDVFSGYGAATPDENY-----APK----- 336

RESULT 13
US-09-844-281-1
; Sequence 1, Application US/09844281
; Patent No. US20020082386A1
; GENERAL INFORMATION:
; APPLICANT: Mangold, Beverly L.
; APPLICANT: Aldrich, Jennifer L.
; APPLICANT: O'Brien, Thomas
; TITLE OF INVENTION: Anthrax Specific Antibodies
; FILE REFERENCE: 38632.0003
; CURRENT APPLICATION NUMBER: US/09/844,281
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,505
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 1
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-844-281-1

Query Match
Best Local Similarity 7.4%; Score 73.5; DB 9; Length 833;
Matches 53; Conservative 29; Mismatches 90; Indels 43; Gaps 10;

QY 8 NLSNLVITDASSLNGVDKLLSAEVEKMLVQK-----APNEGIEVFGLLLYALAA--R 60
```

403 NSQNLVJGEKASLN---KLVATIAGEKVVDPGSGISIKSSNHGIIISVNNVYFAERAGEA 459  
61 TTSPKQORADSDVIFNSFGERNVVVTEGD---LKKVLGDCAPLTRFTNKLRTFGRTFTE 117  
460 TLIKVGDTVQWKFKVTTSDSKLSVVKANPKLVQVQNKTLPTVFT-----TD 369  
118 AYVD---FCIAYKHKLPLQNAAAELG---PAEDSYLAADFGLTCPKLSELQSRKMF- 169  
510 QYGDPFGANTAAIKVELPKTVGVAEGGLDVT--DSSGSGTKVIGVTS--HNVGEGTVHFQ 567  
170 -----ASYALKTEGGVVNTP---VSNLRQLGR 194  
568 NKGATIGLSYNNVNTGKVAFKNFEIVSKVGYGO 602

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2003, 14:44:53 ; Search time 41 Seconds  
(without alignments)  
765.533 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMSDSNLSNLV-TDASSL.....GGVVTPEVSNLRLGRREVX 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 6  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_15Jun03.\*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	198	20	AAW73482 Grapevine leafroll
2	286	28.9	204	16	AAR72682 Sugar beet yellows
3	89.5	9.0	499	23	ABG93273 C. albicans BAX-as
4	86.5	8.7	328	23	ABB91963 Herbicidally activ
5	86	8.7	295	21	AAG20735 Arabidopsis thalia
6	86	8.7	295	21	AAG39167 Arabidopsis thalia
7	86	8.7	360	21	AAG20734 Arabidopsis thalia
8	86	8.7	360	21	AAG39166 Arabidopsis thalia
9	86	8.7	360	23	ABB91962 Herbicidally activ

10	86	8.7	366	21	AAG20733 Arabidopsis thalia
11	86	8.7	366	21	AAG39165 Arabidopsis thalia
12	86	8.7	447	22	AAB96072 Putative P. abyssi
13	85	8.6	599	18	AAW17788 Prage abortive in:
14	84.5	8.5	269	22	AAV97690 GRVAV-5 coat prote
15	81.5	8.2	402	21	AAAB14800 Erysipelothrix rhu
16	81.5	8.2	402	23	ABG07787 E. rhusiopathiae pr
17	81.5	8.2	402	23	ABG08869 Erysipelothrix rhu
18	81.5	8.2	626	21	Erysipelothrix rhu
19	80	8.1	555	23	Lactococcus lactis
20	80	8.1	711	23	Streptococcus poly
21	80	8.1	729	23	Streptococcus poly
22	79.5	8.0	770	18	Rat hemidesmosome
23	79	8.0	263	22	Propionibacterium
24	79	8.0	451	22	Enterococcus faeca
25	79	8.0	477	18	Grapevine leafroll
26	79	8.0	782	22	Xrosophila melanog
27	78.5	7.9	208	23	Lactococcus lactis
28	77	7.8	696	23	Candida albicans e
29	76.5	7.7	200	22	Novel human secret
30	76	7.7	223	20	Grapevine leafroll
31	75.5	7.6	509	15	NADH oxidase. Amp
32	75.5	7.6	606	20	E. rhusiopathiae s
33	75	7.6	358	23	Herbicidally activ
34	75	7.6	404	23	Streptococcus poly
35	75	7.6	883	22	Synechococcus CLPB
36	75	7.6	26926	22	AAU05396
37	75	7.6	31267	24	ABG74786 Human tGSI1 protei
38	74.5	7.5	449	22	Human RGS11 (conne
39	74.5	7.5	1694	21	Beta-TRCP.N/SK22.C
40	74.5	7.5	1725	21	Rat laminin 5 poly
41	74	7.5	359	23	AAAB48460 Rat laminin 5 poly
42	74	7.5	399	21	AAAB43940 Herbicidally activ
43	73.5	7.4	332	21	AAAG10781 Human cancer assoc
44	73.5	7.4	355	21	Arabidopsis thalia
45	73.5	7.4	367	21	Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAW73482  
ID AAW73482 standard; Protein; 198 AA.  
XX AC  
XX AAW73482;  
XX CT 29-MAR-1999 (first entry)  
XX DE Grapevine leafroll virus type 2 coat protein.  
DE GRLAV-2; closterovirus; grape; tobacco; transgenic plant;  
KW disease resistance; virus resistance; beet yellows virus;  
KW tristeza virus; coat protein.  
XX OS Grapevine leafroll virus type 2.  
XX PN WO9853055-A1.  
XX PD 26-NOV-1998.  
XX PF 20-MAY-1998; 98WO-US10313.  
XX PR 20-MAY-1997; 97US-0047194.  
XX (CCRR ) CORNELL RES FOUND INC.  
XX Gonsalves D, Ling K, Zhu H;  
XX WP1; 1999-045307/04.  
XX N-PSDB; AAV08870.  
XX Grapevine leafroll virus (type 2) proteins and polypeptides - and  
PT

PT encoding DNA, useful e.g. to impart grapevine leafroll resistance to  
 PT grape and tobacco plants and detect grapevine leafroll virus  
 XX  
 XX  
 PS Claim 12: Page 44-45; 151pp; English.  
 XX  
 XX This is the amino acid sequence of a 22 kDa coat protein that is  
 CC encoded by open reading frame ORF6 (see AAV08870) of grapevine  
 CC leafroll virus type 2 (GIRAV-2) RNA (see AAV08874). The GIRAV-2  
 CC genome includes 9 open reading frames (see AAV3864-72) for a  
 CC polyprotein, an RNA-dependent RNA polymerase, heat shock proteins,  
 CC coat proteins and proteins of unknown function (see AAW73476-84).  
 CC These can be used to produce antibodies useful for detecting  
 CC GIRAV in samples e.g. by ELISA (claimed). The nucleic acid  
 CC molecules can be used to produce probes and primers for such  
 CC detection, and to transform host cells (especially Agrobacterium  
 CC vitis, Agrobacterium tumefaciens, grape, citrus, beet or tobacco  
 CC cells) and produce transgenic plants (claimed). They can be used  
 CC to impart GIRAV-2 resistance to Vitis scion or rootstock cultivars  
 CC or Nicotiana cultivars (claimed). Because extensive similarity  
 CC exists between hsp70-related sequence regions of GIRAV-2 and other  
 CC closteroviruses, the DNA may also be used to impart beet yellows  
 CC virus resistance to beet cultivars or tritstera virus resistance to  
 CC citrus scion cultivar/rootstock cultivars (claimed).  
 XX

XX  
 SQ Sequence 198 AA;  
 Query Match 100.0%; Score 991; DB 20; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-100;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY : MELMSDSNLSNLVITDASSINGVDKLLSAEVEKMLVQKGNPEGIEVFGLLYALAAAR 60  
 Db 1 MELMSDSNLSNLVITDASSINGVDKLLSAEVEKMLVQKGNPEGIEVFGLLYALAAAR 60  
 QY 61 TTSPKVQRADSDVIFNSFGERNVNVTEGDLKXVLDGCGAPLRTFTNKLRTFGRTTEAYV 120  
 Db 61 TTSPKVQRADSDVIFNSFGERNVNVTEGDLKXVLDGCGAPLRTFTNKLRTFGRTTEAYV 120  
 QY 121 DFCIAYKHKLPLQNLAAELGIPAEEDSVLAADFCTGCPKLSLQOSKMFASMYALKTEGG 180  
 Db 121 DFCIAYKHKLPLQNLAAELGIPAEEDSVLAADFCTGCPKLSLQOSKMFASMYALKTEGG 180  
 QY 181 VNTPTVSNLRQLGRREV 198  
 Db 181 VNTPTVSNLRQLGRREV 198

RESULT 2  
 AAR72682  
 ID AAR72682 standard; Protein: 204 AA.

XX AAR72682;  
 XX AC AAR72682;

DT 25-MAR-2003 (updated)  
 DT 02-NOV-1995 (first entry)

XX Sugar beet yellows virus capsid protein.  
 XX DE DE

XX Sugar beet yellows virus capsid protein; transgenic plant.  
 XX KW KW

XX Sugar beet yellows virus.  
 XX OS OS

XX RU2017820-C1.  
 XX PN PN

XX 15-AUG-1994.  
 XX PJ PJ

XX 27-JUN-1991; 91SU-4950054.  
 XX PF PF

XX 27-JUN-1991; 91SU-4950054.  
 XX PR PR

XX (BIOT=) BIOVACHN INST CO LTD.  
 XX PA PA

XX (IMMU=) IMMUNOBIOTECHN INST.  
 XX PA PA

XX WPI; 2002-667002/71.  
 XX DR DR

P: Arganovskii AA, Boiko VP, Karasev AV;  
 XX WP1; 1995-113715/15.  
 DR N-PSDB; AAQ87853.  
 XX  
 XX Sugar beet yellows virus cDNA fragment encoding capsid protein -  
 PT useful for production of virus-resistant transgenic plants  
 XX  
 XX Claim 1: Column 7-10; 5pp; Russian.  
 XX  
 CC The amino acid sequence of the novel sugar beet yellows virus (SBV)  
 CC capsid protein. The protein has mol. wt. 22.2 kD. The corresponding  
 CC gene was obtained from reverse transcribed RNA isolated from purified  
 CC SBV. The fragment is useful for the production of virus resistant  
 CC transgenic plants by genetic engineering methods.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 204 AA;

Query Match 28.9%; Score 286; DB 16; Length 204;  
 Best Local Similarity 37.6%; Pred. No. 6.2e-23;  
 Matches 71; Conservative 29; Mismatches 87; Indels 2; Gaps 2;

QY 7 SNLSNLVITDASSINGVDKLLSAEVEKMLVQKGNPEGIEVFGLLYALAAATTSFKV 66  
 Db 11 ATFNVSIAQOTCLHGDCDCKLRKNFECLKLGVPEDNLGIALGLCLYSCATIGTSNKV 70  
 QY 67 QRASDVIFNSF-GERNVNVTEGDLKXVLDGCGAPLRTFTNKLRTFGRTTEAYVDFCIA 125  
 Db 71 NVQPTSTFIKASFGGKELVLTGHELSFLGSKLLEKPKKJRCFCRTFKQVYSJRKE 130  
 QY 126 YKHKLPLQNLAAELGIPAEEDSVLAADFCTGCPKLSLQOSKMFASMYALKTEGGVWNT? 185  
 Db 131 YGKLPPIARANRGLPAECHYLAADFISTSTELTDLQOSKLLARENATHTEFS-SESP 189  
 QY 186 VSNLRQLGR 194  
 Db 190 VTSJKQLGR 198

RESULT 3  
 ABG93273  
 ID ABG93273 standard; Protein: 499 AA.

XX ABG93273;  
 XX AC ABG93273;

XX 21-NOV-2002 (first entry)  
 XX DT DT

XX C. albicans BAX-associated protein fragment SEQ ID 504.  
 XX DE DE

XX Bax: Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;  
 XX KW KW

XX vasotropic; vaccine; gene therapy; proliferative disorder; cancer;  
 XX KW KW

XX apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;  
 XX KW KW

XX neurodegeneration; cell death.  
 XX KW KW

XX Candida albicans.  
 XX CS CS

XX WO200264766-A2.  
 XX PN PN

XX 22-AUG-2002.  
 XX PD PD

XX 21-DEC-2001; 2001WO-EP15398.  
 XX PF PF

XX 22-DEC-2000; 2000EP-C870318.  
 XX PR PR

XX 04-JAN-2001; 2001EP-C870002.  
 XX PR PR

XX 09-JAN-2001; 2001EP-C870003.  
 XX PR PR

XX (CANC) JANSSEN PHARM NV.  
 XX PA PA

XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RC;  
 XX P: Contreras RH, Eberhardt I, Luyten WHML, Reekmans RC;  
 XX DR DR



PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137526.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 36-OCT-1999; 99US-0157865.

```

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      8.7%; Score 86; DB 21; Length 295;
Best Local Similarity 26.5%; Pred No. 0.94;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 3;

QY 7 SNLSNLVITDASSLNGV--DKKLSAEVEKMLVQKGAENEGI----EVLGGLLYAAR 60
DE 72 SEQPNFKSYIARLKGIVGDKAMEIINNAFVVSAGPNDILNVEIPSRRLVYP--ISG 13:
QY 61 TTSPKVRQADESDVIFSNFSGERNVVTGDLKKVLDGCAELTRTNKLRITGRTFTAYV 120
DE 132 YGDFILKRLNFEVRELYSLGVNRNLV--GGLPFM--GCLPI-HMTAKFRNIFRCLEHHN 186
QY 121 DFCIAYKHLPCLNAAELGIPAEBSYLAADFLGTCPKLSELQCSRKMFASVYALK-TEG 179
DE 187 KDSVLYNEKLQNLPLQIEASLPG-SKFLYADVYN--PYMEMIQNPSK-----YGFKEIKR 238
QY 180 GVNT 184
DE 239 GCGGT 243

RESULT 6
AAG39167
XX AC AAG39167;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 48421.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX
```

```

PR 25-FEB-2003; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121625.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125786.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 31-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 26-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
```



```
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144894.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147322.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151365.
PR 27-AUG-1999; 99US-0151366.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.7%; Score 86; DB 21; Length 295;
Best Local Similarity 26.5%; Pred. No. 3.94;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLVITDASSLNGV--DKKLLSAEVEKMLVQKGAPNEGJ---EVVFGLLLYALAAAR 60
DB 72 SEQPMKFSYIARLKGIVGCKKAMEIINNAFVVSAGNDFILNYVEIPSRRLXPFIISG 131
QY 61 TTSKPVQADSDVIFNSFGERNVVVTGELKKVLDGCAPIRTFNKLRITFGRTFTTAYV 120
DB 132 YQDFTLKRLNFVRELYSLGVRNLV--GGUPPW--GCLPI--HMTAKFRNIFRCLEHHN 186
QY 121 DECIAYKHLQOLNNAEELGTPAEDSYLAADPLGTCPLKSELQOSRKMFSAMVYALK-TEG 179
DB 187 KDSVLNFKLONLQFIEASLPG-SKFLYADVYN--PMMEMIQNPSSK-----YGFKEKTR 238
QY 180 GVWNT 184
DB 239 GCCGT 243

RESULT 7
AAG20734
-D AAG20734 standard; Protein; 360 AA.
XX
AC AAG20734;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23036.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
```

XX  
PN EP1033405-A2.  
XX  
XX  
PD  
PF  
XX  
XX  
XX  
PR 25-FEB-1999; 99US-0141842.  
PR 05-MAR-1999; 99US-0142154.  
PR 09-MAR-1999; 99US-0142055.  
PR 23-MAR-1999; 99US-014239C.  
PR 25-MAR-1999; 99US-0142803.  
PR 29-MAR-1999; 99US-0142920.  
PR 01-APR-1999; 99US-0142977.  
PR 06-APR-1999; 99US-0143542.  
PR 16-APR-1999; 99US-0143624.  
PR 19-APR-1999; 99US-0144005.  
PR 21-APR-1999; 99US-0144085.  
PR 23-APR-1999; 99US-0144325.  
PR 28-APR-1999; 99US-0144332.  
PR 30-APR-1999; 99US-0144333.  
PR 04-MAY-1999; 99US-0144334.  
PR 05-MAY-1999; 99US-0144335.  
PR 06-MAY-1999; 99US-0144352.  
PR 07-MAY-1999; 99US-0144684.  
PR 11-MAY-1999; 99US-0144814.  
PR 14-MAY-1999; 99US-0145086.  
PR 14-MAY-1999; 99US-0145088.  
PR 14-MAY-1999; 99US-0145085.  
PR 18-MAY-1999; 99US-0145087.  
PR 19-MAY-1999; 99US-0145089.  
PR 20-MAY-1999; 99US-0145192.  
PR 21-MAY-1999; 99US-0145145.  
PR 21-MAY-1999; 99US-0145218.  
PR 21-MAY-1999; 99US-0145224.  
PR 21-MAY-1999; 99US-0145276.  
PR 21-MAY-1999; 99US-0145913.  
PR 21-MAY-1999; 99US-0145919.  
PR 21-MAY-1999; 99US-0145951.  
PR 21-MAY-1999; 99US-0146386.  
PR 21-MAY-1999; 99US-0146388.  
PR 21-MAY-1999; 99US-0146389.  
PR 25-MAY-1999; 99US-0147038.  
PR 27-MAY-1999; 99US-0147204.  
PR 28-MAY-1999; 99US-0147302.  
PR 01-JUN-1999; 99US-0147192.  
PR 03-JUN-1999; 99US-0147260.  
PR 04-JUN-1999; 99US-0147303.  
PR 07-JUN-1999; 99US-0147416.  
PR 08-JUN-1999; 99US-0147493.  
PR 10-JUN-1999; 99US-0147935.  
PR 10-JUN-1999; 99US-0148171.  
PR 14-JUN-1999; 99US-0148319.  
PR 16-JUN-1999; 99US-0148341.  
PR 17-JUN-1999; 99US-0148565.  
PR 18-JUN-1999; 99US-0148684.  
PR 18-JUN-1999; 99US-0149368.  
PR 18-JUN-1999; 99US-0149175.  
PR 18-JUN-1999; 99US-0149426.  
PR 18-JUN-1999; 99US-0149722.  
PR 18-JUN-1999; 99US-0149723.  
PR 18-JUN-1999; 99US-0149929.  
PR 18-JUN-1999; 99US-0149902.  
PR 18-JUN-1999; 99US-0149930.  
PR 18-JUN-1999; 99US-0150566.  
PR 18-JUN-1999; 99US-0150884.  
PR 18-JUN-1999; 99US-0151065.  
PR 18-JUN-1999; 99US-0151066.  
PR 21-JUN-1999; 99US-0151066.  
PR 22-JUN-1999; 99US-0151080.  
PR 23-JUN-1999; 99US-0151303.  
PR 23-JUN-1999; 99US-0151303.  
PR 24-JUN-1999; 99US-0151930.  
PR 24-JUN-1999; 99US-0152363.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-014239C.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 20-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.



```

PR 23-JUN-1999; 99JUS-0140354;
PR 24-JUN-1999; 99JUS-0140695;
PR 28-JUN-1999; 99JUS-0140823;
PR 29-JUN-1999; 99JUS-0140991;
PR 30-JUN-1999; 99JUS-0141287;
PR 01-JUL-1999; 99JUS-0141842;
PR 01-JUL-1999; 99JUS-0142154;
PR 02-JUL-1999; 99JUS-0142555;
PR 06-JUL-1999; 99JUS-0142390;
PR 08-JUL-1999; 99JUS-0142803;
PR 09-JUL-1999; 99JUS-0142920;
PR 12-JUL-1999; 99JUS-0142977;
PR 13-JUL-1999; 99JUS-0143542;
PR 14-JUL-1999; 99JUS-0143624;
PR 15-JUL-1999; 99JUS-0144005;
PR 16-JUL-1999; 99JUS-0144085;
PR 16-JUL-1999; 99JUS-0144086;
PR 19-JUL-1999; 99JUS-0144325;
PR 19-JUL-1999; 99JUS-0144331;
PR 19-JUL-1999; 99JUS-0144332;
PR 19-JUL-1999; 99JUS-0144333;
PR 19-JUL-1999; 99JUS-0144334;
PR 19-JUL-1999; 99JUS-0144335;
PR 20-JUL-1999; 99JUS-0144352;
PR 20-JUL-1999; 99JUS-0144632;
PR 20-JUL-1999; 99JUS-0144884;
PR 21-JUL-1999; 99JUS-0144814;
PR 21-JUL-1999; 99JUS-0145086;
PR 21-JUL-1999; 99JUS-0145088;
PR 22-JUL-1999; 99JUS-0145085;
PR 22-JUL-1999; 99JUS-0145087;
PR 22-JUL-1999; 99JUS-0145089;
PR 22-JUL-1999; 99JUS-0145192;
PR 23-JUL-1999; 99JUS-0145145;
PR 23-JUL-1999; 99JUS-0145218;
PR 23-JUL-1999; 99JUS-0145224;
PR 26-JUL-1999; 99JUS-0145276;
PR 27-JUL-1999; 99JUS-0145913;
PR 27-JUL-1999; 99JUS-0145918;
PR 27-JUL-1999; 99JUS-0145919;
PR 28-JUL-1999; 99JUS-0145951;
PR 02-AUG-1999; 99JUS-0146386;
PR 02-AUG-1999; 99JUS-0146388;
PR 02-AUG-1999; 99JUS-0146389;
PR 03-AUG-1999; 99JUS-0147038;
PR 04-AUG-1999; 99JUS-0147204;
PR 04-AUG-1999; 99JUS-0147302;
PR 05-AUG-1999; 99JUS-0147192;
PR 06-AUG-1999; 99JUS-0147260;
PR 06-AUG-1999; 99JUS-0147303;
PR 09-AUG-1999; 99JUS-0147416;
PR 09-AUG-1999; 99JUS-0147493;
PR 09-AUG-1999; 99JUS-0147935;
PR 10-AUG-1999; 99JUS-0148171;
PR 11-AUG-1999; 99JUS-0148319;
PR 12-AUG-1999; 99JUS-0148341;
PR 13-AUG-1999; 99JUS-0148565;
PR 13-AUG-1999; 99JUS-0148684;
PR 16-AUG-1999; 99JUS-0149368;
PR 17-AUG-1999; 99JUS-0149175;
PR 18-AUG-1999; 99JUS-0149426;
PR 20-AUG-1999; 99JUS-0149722;
PR 20-AUG-1999; 99JUS-0149723;
PR 20-AUG-1999; 99JUS-0149929;
PR 23-AUG-1999; 99JUS-0149902;
PR 23-AUG-1999; 99JUS-0149920;
PR 25-AUG-1999; 99JUS-0150566;
PR 26-AUG-1999; 99JUS-0150884;
PR 27-AUG-1999; 99JUS-0151065;
PR 27-AUG-1999; 99JUS-0151066;
PR 27-AUG-1999; 99JUS-0151080;
PR 30-AUG-1999; 99JUS-0151303;
PR 31-AUG-1999; 99JUS-0151438;

PR 01-SEP-1999; 99JUS-0151930;
PR 07-SEP-1999; 99JUS-0152263;
PR 13-SEP-1999; 99JUS-0153070;
PR 13-SEP-1999; 99JUS-0153758;
PR 15-SEP-1999; 99JUS-0154018;
PR 16-SEP-1999; 99JUS-0154039;
PR 20-SEP-1999; 99JUS-0154779;
PR 22-SEP-1999; 99JUS-0155139;
PR 23-SEP-1999; 99JUS-0155486;
PR 24-SEP-1999; 99JUS-0155659;
PR 28-SEP-1999; 99JUS-0156458;
PR 29-SEP-1999; 99JUS-0156596;
PR 04-OCT-1999; 99JUS-0157117;
PR 05-OCT-1999; 99JUS-0157753;
PR 06-OCT-1999; 99JUS-0157865;
PR 07-OCT-1999; 99JUS-0158029;
PR 08-OCT-1999; 99JUS-0158232;
PR 12-OCT-1999; 99JUS-0158369;
PR 13-OCT-1999; 99JUS-0159293;
PR 13-OCT-1999; 99JUS-0159294;
PR 13-OCT-1999; 99JUS-0159295;
PR 14-OCT-1999; 99JUS-0159329;
PR 14-OCT-1999; 99JUS-0159330;
PR 14-OCT-1999; 99JUS-0159331;
PR 14-OCT-1999; 99JUS-0159637;
PR 14-OCT-1999; 99JUS-0159638;
PR 18-OCT-1999; 99JUS-0159584;
PR 21-OCT-1999; 99JUS-0160741;
PR 21-OCT-1999; 99JUS-0160767;
PR 21-OCT-1999; 99JUS-0160768;
PR 21-OCT-1999; 99JUS-0160770;
PR 21-OCT-1999; 99JUS-0160814;
PR 21-OCT-1999; 99JUS-0160815;
PR 22-OCT-1999; 99JUS-0160980;
PR 22-OCT-1999; 99JUS-0160981;
PR 22-OCT-1999; 99JUS-0160989;
PR 22-OCT-1999; 99JUS-0161404;
PR 25-OCT-1999; 99JUS-0161405;
PR 25-OCT-1999; 99JUS-0161406;
PR 26-OCT-1999; 99JUS-0161359;
PR 26-OCT-1999; 99JUS-0161360;
PR 26-OCT-1999; 99JUS-0161361;
PR 28-OCT-1999; 99JUS-0161920;
PR 28-OCT-1999; 99JUS-0161992;
PR 28-OCT-1999; 99JUS-0161993;
PR 23-OCT-1999; 99JUS-0162142;

Query Match      8.7%; Score 86; DB 21; Length 360;
Best Local Similarity 26.5%; Pred. No. 1.2;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY      7 SNLSNLVITDASSLNGV--DKKLSAEVEKMLVQKGNPNEG---EVVFGLLLYALAAR 60
Db      137 SEQPNFKSVYIARLKGIVGDKKAMEIINKAFVVVSAGPNDFLNYYEIPSRLLEYPFISG 196

QY      61 TTSPKVQRADSDVIFNSFGERNVVVTEGDLKKVLDGCGAPJTRFTNKLRTFTGRTFTYAV 123
Db      197 YQDFILKRLNFVRELYSLGRNVLV--GGLPPM--GCLPI-HM-AKFRNIFRCLEHHN 251

QY      121 DFCIAYKHLPQI-NAAAEIGIPAEUSYLAADFLGTCFKLSE:QQSKKMFASVYALK-TEG 179
Db      252 KDSVLYNEKLQNLPLQIEASLPQ-SKFLYADVYN--PMMEMIONPSK-----YGFKEIKR 303

QY      180 GVYNT 184
Db      304 GCCGT 308

RESULT 9
ABB91962
ID ABB91962 standard; Protein; 360 AA.
XX ABB91962;
AC ABB91962;

```



```
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148566.
PR 16-AUG-1999; 99US-0149366.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 22-OCT-1999; 99US-0160741.
PR 22-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160915.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 8.7%; Score 86; DB 21; Length 366;
Best Local Similarity 26.5%; Pred. NO. 1.3;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLVITDASSLNGV--DKKLLSREVEKMLVOKGAPNEGII----EVVFGULLYALAAR 60
DB 143 SEQPMFKSYIAKLGKIVGDKKAMEIINNAFVVVSAGPNDFILNYYEIPSRRLYEYF--SG 202
QY 61 TTSFKVQRADSDVIFNSFGERNVWVTEGDLKKVLDCAPLTFRTNKLRTFGRTF--EAYV 120
DB 203 YQDFILKRLNFVRELYSLGVRVLV--GGLPPX--GCLPI--HMTAKFRNIFAPCLEHNN 257
QY 121 DFCIAYKHLFQNLNAAAEIGIPAEEDSYLAADFGLGTCPKUSELQOSKRYFASMYALK--TEG 179
DB 259 KDSVLNNEK--QNLLEPQIEASLPG--SKFLYADVYN--PMMEYIQNPCK----YGFNETKR 309
QY 180 GVWNT 184
DB 310 GCCGT 314

RESULT 11
AAG39165
ID AAG39165 standard; Protein; 366 AA.
XX
AC AAG39165;
```



```
PR 27-AUG-1999; 99US-0131065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154775.
PR 22-SEP-1999; 99US-0155135.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0156658.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-016074.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161362.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match      8.7%; Score 86; DB 21; Length 366;
Best Local Similarity 26.5%; Pred. No. 1.3;
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLTVDASSLNGV--DKKLSAEVEKMLVQKGN--EVVFGLLVALAAR 60
DB 143 SEQPNFYSYARLKGIVGDKKAMEIINNAFVVSAGNDFILNYEIPSRLEYPIFS 202
QY 61 T-SPKVRADSDVIFSNFSGERNVVVTEGDKKVLGDCAPLTFRTNKLRTFGRTFT 120
DB 203 YQDFILKRLNFVRELYSLGVNVLV--GGLPPI--GCLPI-HMTAKFNI-FRPLCEHNN 257
QY 121 DFCIAYKHLPOLNAAAE--GIPAEOSYLAADPLGTCPKLSLELQOSRKMFASKYALK-TEG 179
DB 258 KDSVLYNEKLQNLPLQIEASLPG-SKFLYADVYN--PMWEMIQNPSK-----YGFKETKR 309
QY 180 GVNT 184
DB 310 GCCGT 314

RESULT 12
AAB96072
XX AAB96072 standard; Protein; 447 AA.
XX AAB96072;
XX 29-OCT-2001 (first entry)
XX Putative P. abyssi beta-lactamase.
XX Hyperthermophilic archaeon; hyperthermophilic protein.
XX Pyrococcus abyssi.
XX PR2792651-A..
XX 27-OCT-2000.
XX 21-APR-1999; 99PR-0005034.
XX 21-APR-1999; 99PR-0005034.
XX (CNRS) CNRS CENT NAT RECH SCI.
XX (IFRE-) IFREMER INST FR RECH EXPL MER.
XX Forterre P, Thierry JC, Prieur J, Dietrich J, Lecompte O;
XX Querellou J, Weissenbach J, Saurin W, Heilig R;
XX WPI; 2001-126236/14.
XX New nucleotide sequences isolated from Pyrococcus abyssi encode
XX proteins useful in industry -
XX Claim 7; Pages 693-694; 1657pp; French.
XX The present invention relates to the genomic sequence of Pyrococcus
XX abyssi (see AAF86431 and AAF41223-7) and P. abyssi proteins. P. abyssi is
XX a hyperthermophilic archaeon, which is isolated from deep-sea
XX hydrothermal vents. The present sequence is one such P. abyssi protein.
XX The proteins of the present invention have various potential industrial
XX uses, since the proteins are stable at very high temperatures, some up to
XX 110 degrees centigrade.
XX Note: This patent is in the same patent family as WO2000065062, which
XX contains additional sequences as shown in AAB99132-AAB99143,
XX AAH75903-AAH75920 and AAG66436.
XX Sequence 447 AA;

Query Match      8.7%; Score 86; DB 22; Length 447;
Best Local Similarity 26.0%; Pred. No. 1.7;
Matches 40; Conservative 15; Mismatches 41; Indels 58; Gaps 8;

QY 26 KLSAEVEMLVQKGN--EVVFGLLVALAARTSPKVRADSDVIFSNFSGERNV- 84
DB 5 KLSFIVEKVAERK-VPGISII-----KGDVVYAKGFGYRNVE 44
QY 85 -----VVTEGDKKVLGDCAPLTFRTNKLRTFGRTFT 116
DB 45 ARLPSTPTIYIGISTKSFALAIKLVLEGSL--LDD--PVEKFNVIKURPFGEPT 100
QY 117 EAYVDFCIAYKHLPOLNAAAE--GIPAEOSYL 148
DB 101 ---VHLLTHSSGIPSLGYAEAFIDGMVSGDNWL 131

RESULT 13
AAW17788
XX AAW17788 standard; Protein; 599 AA.
XX AAW17788;
XX 19-AUG-1997 (first entry)
XX
```



```

DE Phage abortive infection protein AbiE.
XX
XX
KW AbiE; abortive phage infection protein; phage resistance;
KW PSRQ800; lactic acid bacterium.
XX
XX Lactococcus lactis subsp. lactis isolate W1.
XX WO3720917-A2.
XX
XX 12-JUN-1997.
XX
XX 20-NOV-1996; 96WO-IB01385.
XX
XX 01-DEC-1995; 95US-0565907.
XX (UNIL ) QUEST INT BV.
XX
XX Holier BJ, Kondo JK, Moineau S, Vandenbergh PA;
PI Vedamuthu ER;
XX
XX WPI: 1997-319765/29.
XX N-PSDB; AA768648.
XX
XX Isolated DNA encoding the AbiE protein of Lactococcus - for
PT protecting strains used in production of fermented dairy products
XX
XX Claim 48; Page 30-32; 49pp; English.
XX
XX AbiE (AA017798) is a phage abortive infection protein that increases
CC resistance to phage. It is the expression product of an open
CC reading frame found in the 4.5 kb EcoRI fragment (AA168648) of
CC Lactococcus lactis (L.L.) subsp. lactis W1 plasmid PSRQ800. AbiE
CC acts at, or before, phage replication and has no homology with
CC known Abi proteins. It protects against phages of 936 and P335
CC types, also against c2 when present on a high copy number plasmid.
CC AbiE can be used to impart phage resistance to bacteria, esp. L1.
CC that do not have natural resistance. The recombinant bacteria can
CC be used in the prodn. of fermented dairy products.
XX
XX SQ Sequence 539 AA;
Query Match 8.6%; Score 65; DB 16; Length 599;
Best Local Similarity 24.3%; Pred. No. 3.3;
Matches 36; Conservative 26; Mismatches 46; Indels 40; Gaps 6;
QY 2 ELMSDSNL-----SNLIVTD-----ASSUNGVDK--KLSA 30
DB 261 EFLNEFNLCRENNLIINDNKTVDNPFVDSKSDIFSFFENITSNCKWKEISN 320
QY 31 EVERMLVQKGAPNEG-IEVVFGLLYLAARTTSPKVORADSDVIFNSFGERNVYVTEG 89
DB 321 FIDYCWNEEHLGNKGAIKCFPVI-----TNTLKQKKVETKNI-DNIFSKRMVTFN 372
QY 90 DLKVLDDGCAPLRTFTNKLRTFGRTFE 117
DB 373 VFEKILDLSLKDSRLTNKFLTFEENINE 403
RESULT 14
AAV97690
ID AAV97690 standard; Protein; 269 AA.
XX
XX AAV97690;
AC
XX
XX 08-MAY-2001 (first entry)
XX
XX GLRaV-5 coat protein.
XX
XX GLRaV-5; grapevine leafroll virus; GLRaV infection; GLRaV coat protein;
KW GLRaV HSP70 homologue protein; viral gene mapping;
KW plant disease resistance.
XX
XX Grapevine leafroll virus.
OS

```

---

```

XX WO250105957-A2.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19708.
XX
XX 19-JUL-1999; 99US-0144453.
XX (AGR1-) AGRITOPF INC.
XX
XX Good XC, Monis J;
XX
XX WPI: 2001-147339/15.
XX N-PSDB; AA91259.
XX
XX Novel grapevine leafroll virus polynucleotide useful as diagnostic and
PT probe, for viral gene mapping and for induced plant disease resistance
XX
XX Claim 1; Fig 1; 60pp; English.
XX
XX This sequence represents a grapevine leafroll virus (GLRaV-5) protein
CC sequence of the invention. The DNA sequence can be used in an expression
CC construct. The construct is useful for providing resistance to GLRaV
CC infection in a recombinant plant cell by transforming the plant cell with
CC it, where transcription of the polynucleotide sequence interferes with a
CC normal viral function such as movement, encapsidation or replication of
CC viral RNA. The polynucleotide sequence is expressed as an antisense
CC sequence and encodes a GLRaV coat protein, preferably a defective GLRaV
CC coat protein or a GLRaV HSP70 homologue protein. The GLRaV-5 DNA is
CC useful for the synthesis of GLRaV, as diagnostics and probes, for viral
CC gene mapping and for induced plant disease resistance. It is also useful
CC to detect and quantitate expression of GLRaV in plant tissue prior to use
CC in vegetative propagation, by detecting the presence of GLRaV RNA.
XX
XX SQ Sequence 269 AA;
Query Match 8.5%; Score 84.5; DB 22; Length 269;
Best Local Similarity 24.4%; Pred. No. 1.2;
Matches 32; Conservative 21; Mismatches 57; Indels 21; Gaps 4;
QY 61 TTSKPVQRADSDVIFNSFGERNVYVTEGDLKKVLDGCAPLRTFTNKLRTFGRTFEAYV 120
DB 127 STSPKVSSTNNKTI-TGYDQKEVTVAHDEIKTALDINSIGSFYENTPQFGRAFTAAIV 185
QY 121 DFCIAVKHKLPLNAAAEELGPAE-----DSVLAADFLGTCFCKSEL 162
DB 186 QGSSGKLEV-NTKICASHGVPPNYPYSPDCLHVDARLFGYDASLAAE-LGKMVAINKP 243
QY 163 QGSRKXVFASMY 173
DB 244 SNSNRATHNLY 254
RESULT 15
AAB14800
ID AAB14800 standard; Protein; 402 AA.
XX
XX AAB14800;
AC
XX
XX 11-DEC-2000 (first entry)
XX
XX Erysipelothrix rhusiopathiae protective polypeptide antigen fragment.
XX
XX Protective polypeptide antigen; PPA; Fujisawa strain; swine erysipelas;
KW recombinant production; immunogenic; subunit vaccine;
KW transmucous administration.
XX
XX Erysipelothrix rhusiopathiae.
OS
XX
XX JP3072345-B1.
XX

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - nucleic search, using frame\_plus\_p2n model  
Run on: November 7, 2003, 20:29:33 : Search time 35:0 Seconds  
(without alignments)  
2307.723 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMDSNLSNLVITASSL.....GGVATPVSNLRQLGREVM 198

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DB=xih  
-Q=/cgn2\_1/USPTO\_spool/US09613486/runat\_07112003\_120411\_27095/app\_query.fasta\_1.391  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LCOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09613486 @CNC 1.1 3508 @runat\_07112003\_120411\_27095 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=106 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_cm.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_to.\*  
27: em\_ste.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rnd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	991	100.0	597	6	AR138306 Sequence
2	991	100.0	15000	14	AF039204 Grapevine
3	991	100.0	15500	6	AR138299 Sequence
4	984	99.3	8590	14	GLAV4131
5	921	92.9	18527	14	AF314061 Grapevine
6	297	30.0	15468	14	AF190581 Beet yell
7	296	29.9	612	14	BYU7-295
8	295	29.8	15468	14	AF056575 Beet yell
9	293	29.6	2724	14	BYVCPG
10	293	29.6	5980	14	BYVMBP
11	286	28.9	6746	14	GLBYV3PH
12	286	28.9	15480	14	BYVCA
13	285	28.8	783	6	A41914
14	265.5	26.9	10545	14	BYU51931
15	130.5	13.2	672	14	AF501867 Citrus tr
16	129.5	13.1	672	14	AF184115 Citrus tr
17	125.5	12.7	669	14	AF342894 Citrus tr
18	125.5	12.7	672	14	AF184114 Citrus tr
19	125.5	12.7	672	14	AF342890 Citrus tr
20	125	12.6	1080	14	CTVCOAT
21	125	12.6	19293	14	AY170468 Citrus tris
22	125	12.6	19296	14	CTU16304
23	124.5	12.6	672	14	AF184117 Citrus tr
24	123.5	12.5	672	14	AF184116 Citrus tr
25	123.5	12.5	672	14	AF220502 Citrus tr
26	123.5	12.5	672	14	AF220504 Citrus tr
27	123	12.4	636	14	CTR297702
28	123	12.4	672	14	AF220503 Citrus tr
29	120.5	12.2	589	14	AF456771 Citrus tr
30	120.5	12.2	669	14	AF342895 Citrus tr
31	120.5	12.2	672	14	AF184118 Citrus tr
32	120.5	12.2	672	14	AF342891 Citrus tr
33	120.5	12.2	672	14	AF342892 Citrus tr
34	120.5	12.2	672	14	AF342893 Citrus tr
35	120.5	12.2	672	14	CTVCOATS
36	120.5	12.2	19226	14	CTU56902
37	120.5	12.2	19259	14	AF260651 Citrus tris
38	120.5	12.2	19259	14	CTV18420
39	118.5	12.0	672	14	AF501869 Citrus tr
40	117.5	11.9	587	14	AF220505 Citrus tr
41	117.5	11.9	672	14	AF339088 Citrus tr
42	116	11.7	672	14	AF184113 Citrus tr
43	115.5	11.7	589	14	AF456773 Citrus tr
44	115.5	11.7	672	14	AF501868 Citrus tr
45	114.5	11.6	672	14	AY190048 Citrus tr

ALIGNMENTS

RESULT 1

[illegible]

```

GLNVAEPEYVMSGLALFEIHSKLCITYDCLRIVNFDRSVAPCSDEAQLYVVKQSGTI
VQKEVRLH:KPHEDHDFDGDKISINKARGGNVLYHNLAFNLAKNLEAFPPSRS
FVT:NSVD:LLYEAPGGGKTKTIDSLFLVFKGGEVSMILITANKSQVEILKXVE
KEYNSICQKRPKSG:YITDILMEHRCGDADVLIFDCEFMVHAGSVLACIEF
TRCHKMIFDQKQ:HYIENELCKJYGLDRFVQLQCRVYGNISYPCPQVDCAWLS
TVGNLJATVKGSGEKGSMRINEINSVDDLVPDVGSTFLCMLQSKLBIKSHFIRKG
LTKNLVLTVEAOGELYARVNLVRLAFQDEDEPKSIRHITVALSRHDSLT:YVLAAR
RGATCDALQKAEELNKNKRFVETSPGSGVINLWKVDVEDSRCSASSAPLSVNDP
LNEVPGTAVIDFGDLSADAFSTPFEGGASG:VVRDNISSNITDHCKQRV
<7422..8801
/feature="replication"
/notes="p63; 52 kDa; similar to RNA polymerases of other
c.isteroviruses; presumably expressed via +1 ribosomal
frameshift"
/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="AAC40856.1"
/db_xref="GI:312391"
/translation="SVVRSQAIPIRRKPSIQENLYSPEARVNFSTCDRNTSASMFOEA
YANCLRRCCDLDAFSLRDVLIS:RSIGIECHLEKRTSQ:KALMKDYESPLEIDDE
-CRFLKLVKRDADKVLKLDSSCLTHSAQNIMFRKKS:INALFSPIFNEVKRIMCLKP
NIRKFTMTNRDFAVSVNSVGLQDDVTHIGEDFSKYDQSAFVAFSEVWKELGV
DELLAIMMCGERLSANTLDGQLSFTIENQRKSGASNTWIGNSLVTGLISLYYDVR
NFALYISGDSLIIFSRSE:SNVADICTDMGFETKPSVPYFCXKFWCMGKHFT
FVDPVKLPVKLGAVKEDVSMDFLFTFTSFKDLTSDENDERLOKLAELVALKYEVQ
TGNTLALS:VHCLRSNLSFSLKLYPRVKGMQVYTSVKKALKKSCSLFDSY:YFPG
QAVMWDDE"
8805..9035
/notes="6 kDa; probably membrane-associated; similar to
small hydrophobic proteins of other closteroviruses"
/codon_start=1
/product="putative transmembrane small hydrophobic
protein"
/protein_id="AAC40857.1"
/db_xref="GI:3123912"
/translation="MNQVQFECFLNLAVFAVTFIFILLVFRV:KSPRQKGHEAPV
PVVRGGFSTVV"
9051..10850
/notes="p65; HSP70; similar to heat shock 70 proteins;
identified by sequence comparison"
/codon_start=1
/product="65 kDa chaperone protein"
/protein_id="AAC40858.1"
/db_xref="GI:3123913"
/translation="MVVFGIDFGTFTSTVCYKQGVFSPKQNN:SAVITLYLFSDS
NHTFGYEAESLSNKLKVSQSFYDLKRWKVCSSNLDAYDLRLKHYSVRLVK:QSG
LNETVSGNFGVTSAAHLPGL:ALPIKAVIGCAEGAFACCTGVCISVFANVDSVQ
RNFTDQCVS:SGYQCVYMINPEPSAALASCNISGKKSANLAVYDFGGSFDES:SYR
NN:FVVAASGDLNIGRDVRAFLTHLSLTSJEPDLTLDISN:KESLSKTDAEIVY
TLRGVQGRKEDVRNKHILSVMLPYVNRITKILESTLYAKSMNESARVKDGLVL:
GGSYLPLG:ADVLTKHQSVRILRVSPRAAVAGALYSCLSGGGLLIDCAAH:
VALADRSCHOI:CAPAGAPIPFGSMPLY:ARVKNKSOREVAVFEGEYVKCPNKRIC
GANT:RFDIQVGTGYSAPITFYMDFSISSVGVSFVVRGEGKQVSLTGTTPAYNFSSV
ALGSRVRELRH:SLNKFVFLGLLHRRKDRRLIF:KDEAIRYADS:DIADVLKEYSY
AASALPFDEVDLILGVSQVLRGSLSE:PL"
10777..12432
/notes="p63; putative heat shock protein 90 homolog"
/codon_start=1
/product="63 kDa protein"
/protein_id="AAC40859.1"
/db_xref="GI:3123914"
/translation="MSNYSMESLIFKFPYGEADWKVLSR:IAHSSSEIKTLPD:RLYG
GRVVKGEFSA:LNPSFEQELGLFILSEREVGSKLGIITVBEAAVDLNP:KAYKTA
ETCSPDVKGEQKYSMEDVMNFMPLSLNDVNDKVLTEQCKLSNSGGE:INPDKGRF
VALTFKRDJATD:GAANVRGVDYLIVYMSLFEQTKQSQSGNISLTKYCEYRT
YLGSDULFTRAPDIP:LTGLIDYCKEYVYSSYKRVNFRFFLAMP:SDVF
FVQWKAPDVRLLFASAELELVEPTLSLDSQVVGHLRYVESY:SDPAIDALE
DKLZAILKSNPRLSTLAQWLFYCYGEFRTAQSVVORPQVYKTPDSVGGSEYKVK
DVSEKFFKLQRELNPSLRQFNARAEAFKIPNQN:SFRP:SLKVPREEWYK:
DYFRHANESGI:TBEEILILNINSVDVK:CAERACNLTLP:SAKBFKSNHKNISQSSQE
RRIKDPLVLKDTLYEFQHKRAGSGRSTDLGSRADHARGSG"
12344..13015
/notes="p25; CPd; coat protein duplicate"

```

```

/codon_start=1
/product="25 kDa diverged coat protein"
/protein_id="AAC40860.1"
/db_xref="GI:3123915"
/translation="MSSNT:SVPGVGLJEALETSGVLTTRKEAVKFFNELKNENYSV
DSSRLSDSEVKEVLEKSEKSELASTDDBHFVYH:IFFLIRCAKISTSEKVKYVGS
TYVDGKTYYTLVDWVFNWMSLTKYKRVNGLRAPCCACEDLVLTVPIMSEFKTK
AVGMKGPIVQKEYLGADFLSGTSKLSMSCHBRAVSIIVAAKNVDRASAF:TGGERKIVSLY
DLGRY"
13080..13680
/notes="p22"
/evidence="experimental"
/product="22 kDa coat protein"
/protein_id="AAC40861.1"
/db_xref="GI:3123916"
/translation="MELMS:SNLSNLVITDASSINGVDKLLSAEVEKMLVQKGAPNE
G:EYVFGLLYALAAART:SPKVRADSV:FNSFGERNVVVTGDKKVLIDGCAPLT
RFTNKLRTFGRTFTEAVVDFCIAVKHKLPO:NAAE:SIGIPAEDSYLAADFLGTCPKLS
ELQGRKMFASMYALKTEGGVNT:PVSNLRQLGRVYM"
13680..14165
/codon_start=1
/product="19 kDa protein"
/protein_id="AAC40862.1"
/db_xref="GI:3123917"
/translation="MEDYEKSESLLIRTN:NTMLLVKSDASVELPKLLICSLYV
SGREVC:CCNREELTRDFEGNHHTR:OYDSASAFEFNNSDCVVKFFJSTGSV
FWFLRSETKGRAVRL:RTFPEANNFFPGSHCGTMEYCKQVLTETES:IDSFCERN
R"
14167..14784
/notes="p24"
/codon_start=1
/product="24 kDa protein"
/protein_id="AAC40863.1"
/db_xref="GI:3123918"
/translation="MRVIVSPYEAEDILKRSTDMRLNIDSGVLSKCEIKAFSTITRD
LHCASVOMGVDTGLYORNCAEKRLIDTVESNIRLAOPLVREKVAHVHFKDPEKELV
AFITRKVYGLGVREAVKZEMBSLAKTKVLNKM:LEMAFVMSPRANKMAEWLELKF
PVKFRDILLVDEYELNELCAEDDHHVKNVNGENHDELDQDEG"
14785..15000
/notes="similar to 3'UTRs of other monopartite
closteroviruses"
BASE COUNT 3877 a 3050 c 3820 g 4253 t
OR:GIN
Alignment Scores:
Pred. No.: 1.37e-95 Length: 15000
Score: 991.00 Matches: 198
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Idrels: 0
DB: 14 Gaps: 0
US-C9-613-486-15 (1-198) x APO39204 (1-15000)
QY 1 MatGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db 13084 ATGGAGTGTGATGTCGACGACGACCTTAGCAACCTGGTGATACCGACGCTCTAGTCTTA 13143
QY 21 AenglyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlyLysGly 40
Db 13144 AATGGTGTGACAGAGAGCTTTATCTCTCAAGTTGAAAAAATGTTGGTGCAGAAAGG 13203
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg 60
Db 13204 GCTCTTAAACGAGGCTATAGAGATGCTGCTTACTCTTACTCTTACGACTCGCGCAAGA 13263
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 13264 ACCACGTCTCTCAAGGTTACGCGCGACGATTCAGACGCTATATTTTCAATAGTTTCGGA 13323
QY 81 GluArgAsnValValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100

```

```
Db 13324 GAGAGAACTGGTAGTAACAGAGGGTGACCTTAAGAAGGTACTCGACGGGTGTGCGGCT 13383
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyVal 120
Db 13384 CTCACCTAGGTTCACCTAATAAAGCTTAGAACGCTTCGGTGGTACTTTCTACCTGAGGCTTACGTT 13943
QY 121 AspPheCysIleAlaTyValLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
Db 13944 GACTTTTGATCGGTATAGACAAATTAACCCCACTCAACCGCGCGGGAATGGGG 14003
QY 141 IleProAlaGluAspSerTyLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
Db 13504 ATTCGAGCTGAAGATTCGTACTTAGCTGCAGATTTCTGGGTACTTTGCCCGAAGCTCTCT 13563
QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyValAlaLeuLysThrGluGly 180
Db 13564 GAATTACAGCAAAAGTAGGAAGATGTTCCGCGAGTATGTACGCTCTAAAACTGAAGGTGA 13623
QY 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgGluValMet 198
Db 13624 GTGGTAAATACACCACTGAGCAATCTCGCGTACGCTAGGTAGAGGGAAGTTATG 13677

RESULT 3
LOCUS AR138299
DEFINITION Sequence 1 from patent US 6,197,948.
ACCESSION AR138299
VERSION AR138299.1
KEYWORDS Grapevine leafroll virus 2
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 15500)
AUTHORS Zhu, H.-Y., Ling, K.-S. and Gonsalves, D.
TITLE Grapevine leafroll virus (type 2) proteins and their uses
JOURNAL Patent: US 6,197,948-A 1 36-MAR-2001
FEATURES
    source 1..15500
    /organism="unknown"

BASE COUNT 3984 a 3178 c 3970 g 4368 t
ORIGIN

Alignment Scores:
Pred. No.: 1,44e-95 Length: 15500
Score: 991.00 Matches: 198
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-613-486-15 (1-198) x AR138299 (1-15500)

Cy 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db 13594 ATGGAGTTGATGTCGACACCAACCTTAGCAACCTGTGTATACCGACGCTCTAGTCTA 13643
QY 21 AsnGlyValAspLysLysLeuLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
Db 13644 AATGGTGTCCACCAAGAAGCTTTATCTGCTGAAAGTTGAAAAGATGTTGTGCGAAGGG 13703
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyValAlaLeuAlaArg 60
Db 13704 GCTCCTAACGAGGTATAGAAGTGGTGTTCGGTCTACTCTTTACGCACTCGCGGCAAGA 13763
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 13764 ACCAGCTCTCCTAAGGTTACGCGCGGAGATTCAGACGCTTATATTTTCAATAGTTTCGGA 13923
QY 81 GluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100
Db 13824 GAGAGGAATGTGGTAGTACAGAGGGTGACCTTAAGAAGGTACTTCACCGGGGTGCGGCT 13883
```

```
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyVal 120
Db 13884 CTCACCTAGGTTCACCTAATAAAGCTTAGAACGCTTCGGTGGTACTTTCTACCTGAGGCTTACGTT 13943
QY 121 AspPheCysIleAlaTyValLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
Db 13944 GACTTTTGATCGGTATAGACAAATTAACCCCACTCAACCGCGCGGGAATGGGG 14003
QY 141 IleProAlaGluAspSerTyLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
Db 14004 ATTCGAGCTGAAGATTCGTACTTAGCTGCAGATTTCTGGGTACTTTGCCCGAAGCTCTCT 14063
QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyValAlaLeuLysThrGluGly 180
Db 14064 GAATTACAGCAAAAGTAGGAAGATGTTCCGCGAGTATGTACGCTCTAAAACTGAAGGTGA 14123
QY 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgGluValMet 198
Db 14124 GTGGTAAATACACCACTGAGCAATCTCGCGTACGCTAGGTAGAGGGAAGTTATG 14177

RESULT 4
LOCUS GLAV4131
DEFINITION Grapevine leafroll-associated virus 2 genes encoding RNA polymerase and coat protein, hsp70, hsp90 gene and ORF2, ORF7 and CRF8.
ACCESSION Y14131
VERSION Y14131.1
KEYWORDS coat protein; heat shock protein 70-like; ORF1; ORF2; CRF3; ORF4; CRF5; CRF6; ORF7; ORF8; RNA polymerase.
SOURCE Grapevine leafroll-associated virus 2
ORGANISM Grapevine leafroll-associated virus 2
REFERENCE 1. Closteroviridae; Closterovirus.
AUTHORS Abou-Ghanem, N.
TITLE The nucleotide sequence of the 3' terminal region of grapevine leafroll associated closterovirus 2
JOURNAL Unpublished
AUTHORS Abou-Ghanem, N.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1997); N. Abou-Ghanem, Universitat degli Studi di Bari, Dipartimento di Protezione delle Piante and Centro Studio virus e virosi delle colture mediterranee, via Amendola 165/A, I 70126 Bari, ITALY
REMARK revised by [3]
REFERENCE 3 (bases 1 to 8590)
AUTHORS Abou-Ghanem, N.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-1998); N. Abou-Ghanem, Universitat degli Studi di Bari, Dipartimento di Protezione delle Piante and Centro Studio virus e virosi delle colture mediterranee, via Amendola 165/A, I 70126 Bari, ITALY
COMMENT On Jan 10, 1998 this sequence version replaced gi:2369864.
FEATURES
    Location/Qualifiers
        1..8590
            /organism="Grapevine leafroll-associated virus 2"
            /mol_type="genomic RNA"
            /db_xref="taxon:64003"
            /lab_host="N.benthamiana"
            1165..2427
                /note="ORF1"
                /codon_start=1
                /product="RNA polymerase"
                /protein_id="CAA74561.1"
                /db_xref="GI:2369865"
                /db_xref="SPTREMBL:O39852"
                /translation="MFGEAMMNCRLRCFCDFDAFSSLRNDVISTRSIGIEQWLKGRTPSRIKALMKDVESSPLEIDDEICRFKLMVKRDARKVLDSSCLTKHSAQNIHFHRSKINAIFSPIFNEKNRIMCCLKPNIKFTFTEMTNRDFAVSVMGLGDDVDVYHIGVDFSKYDKSQDAVKAFAFEVMYKELGVDEBELLAIWMCGERLSIANTLDGOLFNIENQKSGASNTWIGNSLVTLGLSLYIDFRNFPEALYISGDSLIFSRSEISNYADDICTDNGFETKFMSPVPFCFSKFWMCVGHKHTFFVLDPYKLFVKLGAVKEDVSMDSLFFETFTSPKDLTSDFN"

CDS
```

```
CDS
DELIQKLAELVAKYEVQNTGNTTLASVHICLRSLNFSKSLYPRVKGQVFTYSVK
KALLKSCGSLPDSFV:PFQAVMWDE"
2491..2661
/notes="ORF2"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAA74562.1"
/db_xref="GI:2369866"
/db_xref="SPTREMBL:O39853"
/translation="MNQVLOFELFLNLAVFAVTF:FL:VPRVKSFRQKGHEAPV
PVRGGGFSVW"
2677..4476
/gene="hsp70"
2677..4476
/gene="hsp70"
/codon_start=1
/product="heat shock protein 70-like"
/protein_id="CAA74563.1"
/db_xref="GI:2369867"
/db_xref="SPTREMBL:O39854"
/translation="MVVFLDGFDTFTSTVCVYKGRVFSEKQNSAYIPTLYL:ESDS
NMHTFGYEASLNLKVGFSFYDLKRVGCDSSNLDAYLDB:KPHYSVRLVK:GSG
LNTVSGNPGGVKVS:EAHLFGULIFKAVISCREGAFACCTGVICSPVANYDSQ
RNTFDQVSUSGVCCTVY:NEPSAALSAACNSVGKSKAN:JAVDFGGGTFDVSIIISY
RNTFVRASGGD:NLGGRDVDRAPFLHJFSLTLEPDLT:VYSNLKES:SKTDAEIV
TERGVDRKEDVRVKNILTSMPLIVNRTLKILESTLKSAYKMSNESALVXCDLV:J
GSSYLPGLADVLTHQOSVTRILRVSDPRAAVGALYSSCLSGSGG:LLIDCAHT
VA:ADSSCHOIICAPAGAPIPESSGMPLYLARVKNKSOREIAVFEQVYKCPKNKIC
GAN:RFDGVGTDGAPVTFYMDFSISSVGAVSFVVRGPEGQVSLTCT:PAYNFSY
ALGSRVREHISLNNKVFUGLJLHRAKDRRIILFTKDEAIRVADSI:DIADVLKEYSY
AASALPDEDEVLKSGVQKVLGRSRLEEIPL"
4403..6058
/gene="hsp90"
4403..6058
/gene="hsp90"
/codon_start=1
/product="heat shock protein 90-like"
/protein_id="CAA74564.1"
/db_xref="GI:2369868"
/db_xref="SPTREMBL:O39855"
/translation="MNSYSHESL:PKPYGEADHKKLYLSRIAHSSEIKTLPDIRLYG
GRVKSFEPSALPFEQGLGFI:LSEREVGSKLCIGTVEEAYD:JNPKAYKPTA
ETCSKPDVGQGYKQEDVNMVFNRLNDVNDKYLAEQCKSLNSGEL:YNPDCKGRF
VALTFKDRDPTDGAEDVNRVGLDYLIVAMSLEFQRTQKSGSNGISL:YKVCYEVIR
YLSGTLPTFPAPDRIP:PLTGIHIDFCKEYNI:FYSSVKNVDNPRFLAKYMLPLISVF
FQVMKAPADVPRLFLLSAEALT:EYPTLSLIDSVVGHILSVYESY:SDPAIDALE
DKLEAILKSNPRLS:TAQLWVGFCYGEFRTAQSVVRBPVYKTPDSVGSEFINMK
DVEKFKLQREDPVS:LRQCFNGARAHAFAKIFKNGNISFRPISRLNVPREFWYLN:
DYPFRHNRSGJTEBEEILILNINISVDVRKJCAESACSTLPSAKRFSKNHKSNISQSRQE
RRIKPLVLVKDTLYEFQHKRAGMSRSTRDLGSRADHAKGSQ"
5970..6641
/notes="ORF5"
/duplicate_gene"
/codon_start=1
/product="coat protein"
/protein_id="CAA74565.1"
/db_xref="GI:2369869"
/db_xref="SPTREMBL:O39856"
/translation="MSSNTSPVPGGLEALETSGVLLTTRKEAVDKFEFNLKNEYSSV
DSRLSDSEKVEVLEKRSFKSELASTDEHFYHII:FLIRCALKISTSEKRYVGSB
TYVVDKTYVLDAVFNKMSKSLTKKIKRVNGURAFACEDLYLT:VAP:MSERFKT
AVGMKGLPVQKEYLGACF:SGTSKLSMDHRAVSIVAANKVADRSAFTGGERKIVSLY
DLGRY"
6710..7306
/notes="ORF6"
/codon_start=1
/product="coat protein"
/protein_id="CAA74566.1"
/db_xref="GI:2369870"
/db_xref="SPTREMBL:O39857"
/translation="MLMSDSNLSNLVITDASS:NGVDKKLLSAEYVVKLVQKGPAPNE
GIEVFGLLYALAAARTISPKVRADSDVIF:SNSPGERNVVTVEGDLKKVLGCGCAPLT
RFTNKLTFRGTTEAYVDFCIAKYHKHLPQLNNAARELGI:PAEDSYLAADFLGTCPKLS
```

```
CDS
ELQSRKMFASVYALKTEGGVVNT:VSNLRQLGRREVW"
7306..7793
/notes="ORF7"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAA74567.1"
/db_xref="GI:2369871"
/db_xref="SPTREMBL:O39858"
/translation="MEDYEKSESIIILRTNLNTML:JVVKSDASVELPKLII:CGLYLV
SGREVEITCNREELTRDFEGNHHIVIRRI:LOYDSFAFEFNKSDCVVXKFFLETGSV
FWFLESTKGRAVRLHRTFFEANRFFFGSHCGTMEYCLKQVLTETESIIDSPCEERN
R"
7793..8410
/notes="ORF8"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAA74568.1"
/db_xref="GI:2369872"
/db_xref="SPTREMBL:O39859"
/translation="MRVIVSPVEADILKRTDMLRNI:DSGVLSTKECI:KAPSTITRD
LHCASAYQWQVDTGLYQORCAEXRLIDTVESNIRLAQPL:YREKNAVHFCDEPKELV
AFITRKVELTGVGMREARVREKRSLSLTKVUNKMSLEKAVF:YMSKAWNAEWLELKFS
PVKIFRDLJLDVETNELCAEDDVHVQKVNENE:DNHDLLELCQ38C"
BASE COUNT 2344 a 1686 c 2135 g 2425 t
ORIGIN
Alignment Scores:
Pred. No.: 3,738-95 Length: 8590
Score: 984.00 Matches: 197
Percent Similarity: 99.49% Conservative: C
Best Local Similarity: 99.49% Mismatches: 1
Query Match: 99.29% Indels: C
DB: 14 Gaps: 0
US-09-613-486-15 (1-198) x GLAV4131 (1-8590)
QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
DB 67.0 ATGGAGTTGATGTCGACAGCAACCTTAGCAACCTGGTGATACCGACGCTCTAGTCTA 6769
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
DB 6770 AATGGTGTGACAGAGAGCTTTATCTCTCTGAAGTGTGAAAAATGTTGGTCAGAAAGGG 6829
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAla 60
DB 6830 GTCTCTAACGAGGGTATAGAGTGGTGT:CGGTCTACT:CCCTTTAGCAGCTCCGGCAAGA 6889
QY 61 ThrThrSerProLysVal:GlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
DB 6890 ACCACGTCCTCTAGGTT:CAAGCGGCAGATTCAAGCGTTATATTTCAATAGTTCGGA 6949
QY 81 G:uArgAsnValValVal:ThrGluGlyAspLeuLysValLeuAspGlyCysAlaPro 100
DB 6950 GAGAGGAATGGTAGTAGAACAGAGGTGACCTTAAAGAGGTACTCGACGGGTGTGCGGCT 7009
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120
DB 7010 CTCACATAGTTCACTAATAACTTAGAACGTTGGTGGTCTACTTTTCTACGTAGGCTTACGTT 7069
QY 121 AspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaAlaGluLeuGly 140
DB 7070 GACTTTTGTATCGGTATAAGCACAAATTAACCCCAACTCAACGCGCGCGGAATTGGGG 7129
QY 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
DB 7130 ATTCCAGCTGAAGATTCGTACTAGTCAGATTTTCTGGGTACTTGGCCGAAGCTCTCT 7189
QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180
DB 7190 GAATTACAGCAAGTAGGAAAATGTTCCGAGTATGTACGCTCTATAAAACCTGAAGGTGGA 7249
QY 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluValMet 198
```

250 GTGGTAATACACGAGCACTCGGTACAGTACAGTAGGAGGAGGAGTTAG 7303

AF314061 16527 bp ss-RNA linear VRL 06-NOV-2002

DEFINITION Grapevine rootstock stem lesion associated virus methyltransferase/helicase polyprotein, RNA-dependent RNA polymerase, p6, HSP70-like protein, HSP90-like protein, coat protein duplicate, coat protein, p19, and p24 genes, complete cds.

AF314061

AF314061.1 GI:24636914

DESCRIPTION Grapevine rootstock stem lesion associated virus

SOURCE Viruses; SARNA positive-strand associated virus

ORGANISM Closteroviridae; Closterovirus.

REFERENCE 1 (bases 1 to 16527)

AUTHORS Zhang, Y. and Rowhani, A.

TITLE Nucleotide sequence of grapevine rootstock stem lesion associated virus

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 16527)

AUTHORS Zhang, Y. and Rowhani, A.

TITLE Direct Submission

JOURNAL Submitted (16-OCT-2000) Plant Pathology, University of California, One Shields Ave., Davis, CA 95616, USA

FEATURES

source

1..16527 /organism="Grapevine rootstock stem lesion associated virus"

/mol\_type="genomic RNA"

/db\_xref="taxon:167634"

107..8902

/note="327 kDa"

/codon\_start=1

/product="methyltransferase/helicase polyprotein"

/protein\_id="AA063466.1"

/db\_xref="GI:24636915"

/translation="MSSLAISALPSAAVQALSFQGPAAATVAKSFYMTSLPSCTCFQTC

EFSSFLICASFQKKTFLSFSSVREALFSELSTGFCCEFIPLGHEHFLPVP

QCGLPSVCLGVNGSFAVLN:SPFP:ADVASLMDVGVVTPFKVR:SKRAKLNKRAF

STARAAAR:RSKVGNGKEVIRSHVQ:LPAGLRFQNKNGTELIIGRDVGVARVKT

AAPVRRRGGRGVPLIP:VAVSPFGVYRCVAPSGAICAPKNLASRAANDSRAE

KILSARRKLQONFNSVAKAKARVATASRVMAEGEVTPPRHIFGSGSEPPFPSS

SGMARSVJREPAIVT:CGASAPAKIETPVSVKKAD:SPAAVPREALT:PRVAKGEA

IVLSGVVHEVLAIREGULIRPVGGMRESRFESSHGCYKIVANVR:NTSWLACTKV

NEKFSVAIPCSVDYVAMLYVSGGKFPVLXSRCEYDGLCYMAHMYFJCAFPCCP

RESCYPLGWSVARLAKULVRKFGDEALNIGVGYVSRRAFCYDYSKYARS:LMRL

AGYGVGDEVHTSLASTVYLKPEGRANDRPKISIEGSCFCE:RYVESVNPAPR:JRSE

KVLVRKPEVNGVSTESADVGSNVRKRDARIDRFKSHLAGSVNKKR:IAAGV:RYR

VGGMDPNSFLTOAGVHLLVRRTKSSSVCEVELSTPDGLLRDVI:PCSRDVAALSF

AACGRFP:VLMTRDKYKNCYLAHCYASASALLKGFHAPVDCIGANPTAKLGRMV

SULGDRS:SLNLYGSTS:RGI:FHCDYDAAYKDLRPM:SAIVAGKEMGEVAND:TPA

XKQITLHATCRJ:GNSDLSLIEKOD:DFANDVQSLKRDIPRVIRFFVNSVETQ

SLTRFVPELFELSHEDHFAAASRLLENETVRLGCHGSVD:GSCFLPHLSKT

STWHVCPVJGKQARVRRELQYSSVRSVGDEKILEGPNVDVCHYPLGVEHES

QRMVQVQVJAS:LYEICSAWIKKARITY:TWVTCGEF:JDRGECVMES:LDCCIEVD

HADVMYKFGSSCYSHKLSI:IKDITTPY:VLMGLGFLSVEMVEIRMGNYFKITKSEV

CNIVSKILLRYRANSDVUKLIPRDKKRMCLPGYD:VLDCKFVSVPDYVVCN

CSANVSKTEWVNFISKSRVIR:SGKIIHKDNLDLKIVSFSAFVWLASGRSLA

SEYLANLGHFGSCGCFIATFVLREBEIRNMFLNFCJLGLVKRVAFATLDVNSFD

LDATMESITDPAECVAVDLLEELGLLRANAENEKTRNLGDSMAAKLASEI:VVDIEAK

PVAKHAEHPKSEVEKSEKHAFLPGLRGSGKAGVLKEPLDFVLSASRLSCVADQCRIR

SVVHYLDTAVSFVDQNVENLFTMRVLSKGYSCSPYLANRGLANRVASA:VOIVREI

AVLCANVUSKAKVAVTFSAAVCAMFNCSGSGGOREYKSYMHRYTQVLEDTI:FPED

STYLPLELSA:ICGAITVTSFGSTVSUNAFLLQITKGFSLDELVRNVNVRTHSLST

TATDGMVRSVFSQI:SHLLVGNNTLAYQIAFLTVPLLVKKCVGLIFMSDEITYSG

FIKHGVSEFSFITNLKRGQLVDLKVAVQGVDFSNKHLFRDVTQTAIRV:WQIP

TAVVDGLKCAASAVSGVIGVCKTNVGEKSRGGDDHEEYFCDSSETP:RGLTGGSRSG

FSTLFFESFVLRPARIT:IGRI:KHRCFSKIPESISLKIHE:RTPLLGKXPSDVGYVEFL

SEDAVALSISIECQSIENLIGDFRDMADDSLTDMNYDFDSEFLASVNG:PCAGLR

GGGRVGMALAILNVYNSLIMARAASKNSTFTSLIIFRGLVDLRIRNIFSNKSEVSE

LVRYKYVNF:KSMRGLGLRNEHVNSFGELVPLLRNKLSETFIIRAVPMSKWKV

CLSAERERFFALVCLVLTGFPYFARPLGGVCDKLLSLRRAYRVIIVFLKERCDA

MNGLRCLFHDNGDSTAGLKGGSQOTLWGLLLRVLRFALPSLGVGSKCVFSHDL

FF:CVSLVEEVNMIKILGCLSPFSRPISALAFALLOPQRFVYSGKSLTFLPFLK

ISPRVFKHGLFDFGKISVNRNRSGNVLVYHNLFLASNLFLAGYFSRSFVFN

VIFSNVTKDGBERLSEKMGVEALRSLRI:FEELDAVVEKLNKRNEADDSFSP

PMHESHTDVTSPGSKANSVGTALERAVLVEDAVKSEKISMAKDLVRKCTVSDDE:TN

TLSDDEAIEBVPVSEERDSPKTVRVSEVNLNLSNFSFPKPIVDNDKGTGLINAV

RFYMQEALFEIHSKLCAYDQLRIVNPFERSMAPCNEADQLYVRKNGSTIVQGEV

RSNIKPHDHFDFGKISVNRNRSGNVLVYHNLFLASNLFLAGYFSRSFVFN

SVLILYEAEPGGGKTTITLDFLKVKKEGVTMILTANKKSQVAILKAEKVDISNI

ECOKRDKTPKKSIVTIDAYLMHRCGDILFIDECFVMAHGSCLACIEFTRCHKV

MIQDSRQIHYIERNELDKLYGDLRFVYDLOCRVYGVYRCPMVPCANLVSVDNM

IATVKGESGSKSRINEINSVDLVDMGSTYLMQSEKLEISKHFIRKGLSKUNV

LTVHEAGQETARVNLVRKFQDEDFPKSI:RH:TVALSRRHTSLTITVLAARAGATC

DA:OKANELVNRKRVFTFGGGSVINDVKVDVEDNSRCKASAPLUSV:NDF:SDVNP

GTVIDFGDLSADFTGPFECGASGIVVRDNISSNITDHDQKV

<8901..10280

/note="53 kDa; translated via ribosomal frameshift"

/codon\_start=1

/product="RNA-dependent RNA polymerase"

/protein\_id="AA063467.1"

/db\_xref="GI:24636916"

/translation="SAVRSCAIPRRKPSLQENLYSPEARNYNFSCTCDRTFASMFGEA

MANNCIRRFDELTFFSLRDDVISITRSGTEQWLEKRTPSQIKALMDIESP:ODE

ICRPK:MXKRDAAKVDLSCLTKHSPAONIMFHRKSI:NAIFSP:INEVKNR:MSCIKP

NIKFPTMTNRDFAVSVNMLGDDYVHIGVDFSKYDKSDAFVKSFEVNYKELGV

DEELLAIWCGERLSANTLDGOLSFTIENQKSGASNTWIGNSLVTLGILSL:YDVR

HFAVLVSGDSDSLIFSRNKISNYADDICADMGFETKFPSPVPIFCSFVWVGCHKT

FVPDPYKVLKGAKEVDMSDFETFSFKDLTSDFNDE:V:OKLAEVALKYSQVQ

SGNTALALSIV:HCLRSNFI:FSKLYPRVKGWQVYTVSVKALPKSGCSLYSVTPFG

HAVWVWDE"

10373..10543

/note="6 kDa"

/codon\_start=1

/product="p6"

/protein\_id="AA063468.1"

/db\_xref="GI:24636917"

/translation="MQVQLQFECLYLQAVPAITFVFILLIRV:KSFERNVHETPI

SAVGGQGFSTV"

10552..12351

/note="65 kDa"

/codon\_start=1

/product="HSP70-like protein"

/protein\_id="AA063469.1"

/db\_xref="GI:24636918"

/translation="VVVFGLDGCTTSTVCYKDGKQVYKQNSAIVITLYLFSQET

NHWVFGTEASLHMNAKVRGSPFDRLKRVGGSSNFSDFDRLKPHYAVRPYKIG:G

LNDTVIIGNYGTVRSEVHLPLGIASF:IKAMVSCAENAFSCTCTGVCVPEANYDSVQ

RNFTDQCVSLGQCVVMNPEAAALSACNLINKKSANLAVYDFGGCTPDVSIISYR

NTFTVVASGGDLNGLGRDVRTFLTH:FLSLTSLEPLSLD:SNLKESLSKTDAE:VY

TLKSGDDKEDRVNKSILJTVMLPYVNRNTNI:EAATLSYAKNMGSLRVSQCLVLI

GGSSYLPGLDELILSKHRRIDRLKVADPEAAVAGCALYSSCLSSQGLLVDCAHST

IATAGMSDQILICAPAGAPIPFSCTPLPLKARNRSQ:Q:AVFESEYKCPNRKIC

GSNIKEDIGTIVSYSPITFLDVT:ISSVGAISFVVRGPEKGVSLTGTPAIVSSV

VLGTRSVREHIS:SNKVLGLL:HRKADRL:LFVRSGEAVRYVETVEADV:KEPFSY

NASLSLPDEDEPEPLGKGSVKQLRGGRLEBPL"

12278..13882

/note="62 kDa"

/codon\_start=1

/product="HSP90-like protein"

/protein\_id="AA063470.1"

/db\_xref="GI:24636919"

/translation="MSNHSWGSFKKYVGEADWRKYSKSI:VAHTANIKTLPDQVLYG

GRVKSSEFASATPSPEOGLGLMLSERDQWCKMCGITVEAMVQITNP:EAAPLTA

EXCDPDKTEGQKTYTHEDVMNFRNLNDINDKILAEQWSSLSNSGVLINPDKREF

VALTFKNNELVNDSNANIDCRVGYLVVAMS:LFNCRSKQSGAGNLSLVEKYCAYIRI

YLENTDLYFSPDRIPLLTGILDFCKEYNI:FTYTKRNVDFRFFLTNP:PLISDI

FQWPKVAPDVRLLFDISAEALTLEVTLSLDSQVVI:GHVLRFFVESYTAGDALE

EKLDA:MKSNPHLS:TAQWLWVFCYGYGERTAQORVORPGVYKTPNSVGFEINMK

NVTEFPDKIQOREVNPVSLRQFNQGAHAFAKVPKKNISF:KPI:SLNLI:PREFWYLV

DYTFHANSRGSSEBEEVLNNISIVRVKUCAERACSTLPSAKRFSKHKSSVPSLQKE

KNYDPLIALRNSLYEFRRHNRGRV"



```
CDS
13945..14516
/note="25 kDa"
/codon_start=1
/product="coat protein duplicate"
/protein_id="AA63471.1"
/db_xref="GI:24636920"
/translation="MSSGTGIAVGGSEALTSQFVLSTRKZAVDRFFNELRKNYSVV
DSRLSDQVKEVLEKSKFSELASLDEHFVYHIIFLIRSAIISTDKVKYVSSH
GIEVFLGLYALAAATTPKQVADSDIIFONTYCKTVVTEGDLKKVLEGCAPLN
RPTNKLRTGRFTFAVVPFCVAYKHWPCLNAAELGIPAEDSVLAADFLGACPKLS
ELQOSRKMPASMYA.LKTEGVVNTPVSNLRQLGRREV."
15200..15685
/note="19 kDa"
/codon_start=1
/product="coat protein"
/protein_id="AA63472.1"
/db_xref="GI:24636921"
/translation="MELMSEDLNGLHVIITDASLNGVDVKLLSAEVIKLVQKGPQS
GIEVFLGLYALAAATTPKQVADSDIIFONTYCKTVVTEGDLKKVLEGCAPLN
RPTNKLRTGRFTFAVVPFCVAYKHWPCLNAAELGIPAEDSVLAADFLGACPKLS
ELQOSRKMPASMYA.LKTEGVVNTPVSNLRQLGRREV."
15687..16304
/note="24 kDa"
/codon_start=1
/product="p24"
/protein_id="AA63474.1"
/db_xref="GI:24636923"
/translation="MEVIVSPVEAEILTRSTEMLRNIDSGALDTRCICAKFTSLTRD
LHCASQYQWVDITGLCORDCAEKRLIDTVESNLRHAHPLIREKVATHFCKDEPKELY
AFTRKYVXLGTGAAREAVKIEKRLSKVLNKLSEMAFYMSPRANKVAEWLRLRYT
PLKIYRDLGLVDVITLALCEDEIHVVVTKOEDENHDLQEDC"
BASE COUNT 4250 a 3429 c 4224 g 4624 t
ORIGIN
Alignment Scores:
Pred. No.: 5.26e-88 Length: 16527
Score: 921.00 Matches: 180
Percent Similarity: 95.96% Conservative: 10
Best Local Similarity: 90.91% Mismatches: 8
Query Match: 92.94% Indels: 0
DB: 14 Gaps: 0
US-09-613-486-15 (1-198) x AF314061 (1-16527)
QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerLeu 20
Db 14604 ATGGAGTTGATGCCGATGATAACTTGGAGCGCTCGTCATACCGACGCTTCTAGTTG 14663
QY 21 AspGlyValAspIlysLeuLeuSerAlaGluValGluLeuMetLeuValGlnIlysGly 40
Db 14664 AATGGTGTGCATAGAAGAACTGCTGTGCGAGAGTCTATAAATGCTTGTGCAGAGGGG 14723
QY 41 AProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyAlaLeuAlaAsn 60
Db 14724 GCGCCGAGCCAGGATTTGAAACAGCTTTTGGCTTACTGCTTACCGCTTCGAGCGAGA 14783
QY 61 ThrThrSerProIysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 14784 ACCACATCGCCCAAGTTCAGAGGGGTGATTCTGATATATCTTTCAAAACACTTACGGT 14843
QY 81 GluArgAsnValValThrGluGluAspLeuIlysValLeuAspGlyCysAlaPro 100
Db 14844 GACAAGACTGTTGTCGTGACAGAGAGTGATCTCAAGAAAGTGCTTGAAGGGTGCAGCCT 14903
```

```
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyVal 120
Db 14904 CTTAAATAGATTCACTAAACAGCTCGAACTTTCCGCCGCGACGTTTACGGAGCGGTATGTT 14963
QY 121 AspPheCysIleAlaTyLysHisLysLeuProGlnLeuAsnAlaAlaAlaGluLeuGly 140
Db 14964 GATTTCGCGGGTGTACAGCATAGATGCTCAGCTCAACGCGCGCGCAGAAATGGGG 15023
QY 141 LeuProAlaGluAspSerTyLysAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
Db 15024 ATACCAGCAGAGAATTCCTACCTAGCTCGGACATTTTATAGTGTCTTCCCGAAACTCTCT 15083
QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyAlaLeuLysThrGluGlyGly 180
Db 15084 GAATTCACACAAAGTAGAAATGTTTGCAAGTATGTACGCTTTGAAGACTGAGCGCGGA 15143
QY 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgAGGiuValMet 198
Db 15144 GTAGTAAATACGCCCGTGTAGTAATTCGCCCAATTAGGTAGACAGAGATGATG 15197
RESULT 6
AF190581
LOCUS Best yellows virus strain BYV-4, complete genome. linear VRL 04-JAN-2000
DEFINITION
ACCESSION AF190581
VERSION AF190581.1 GI:6492367
KEYWORDS
SOURCE
ORGANISM
Beet yellows virus
Beet yellows virus
Viruses; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
1 (bases 1 to 15468)
Peremyslov,V.V., Hagiwara,Y. and Dolja,V.V.
HSP70 homolog functions in cell-to-cell movement of a plant virus
Proc. Natl. Acad. Sci. U.S.A. 96 (26), 14771-14776 (1999)
20079557
MEDLINE
1061288
2 (bases 1 to 15468)
Peremyslov,V.V., Hagiwara,Y., Alzhanova,Y. and Dolja,V.V.
Direct Submission
Submitted (27-SEP-1999) Botany and Plant Pathology, Oregon State
University, 2082 Cordley Hall, Corvallis, OR 97331, USA
Location/Qualifiers
1..15468
/organism="Beet yellows virus"
/mol_type="genomic RNA"
/strain="BYV-4"
/db_xref="taxon:12161"
/note="Californian strain; RNA transcript derived from
this cDNA clone causes formation of red necrotic lesions
in Claytonia perfoliata and systemic infection in
Nicotiana benthamiana"
Join(108)..7997,7999..9381)
/note="The polyprotein encompasses domains for papain-like
leader proteinase, putative methyltransferase, RNA helicase
and RNA-dependent RNA polymerase (RdRp); ORF1a/b;
papain-like leader proteinase is capable of
autoproteolytic release from the rest of the polyprotein,
and is a replicational enhancer. RdRp is presumably
expressed via a +1 ribosomal frameshift"
/codon_start=1
/product="p348"
/protein_id="AAFI4300.1"
/db_xref="GI:6492368"
/translation="MAFLNVSAPSCAPAFAPAHAGASPIVDPSPFCVFRYSDDIGH
RHLTLDSPVPLFLNARVHLRASTNDPLPLGFLHAETFLVLENGSSAPSPIS
RLIDFVNRPSPVPTPSLVSSTPSLFCDFELCYCKPGCPVEIASPTSP
CLNSVCAQIPTHAEMLISRPPTKLPAGLLQFKKPKRPTLIHESGLAKTS
ALGTSKPNRPITVKASGEKYDAYEISKDFSRRRQOTPRVSHKPKINKAVE
PFFTFEPKDKKRSLPTKDEGFTTFGLRFLSETPEEPPLFRKVEIPVVK
HAPVAVKVPRTFRVATTGAENYVATQCSRPNRHPILRSASYTFGKMPPLQRF
MKKKDYVVKRSKVSSCVTSKPLEALTSLKMLPRYSNLSRLKFDYDHFIGDDFEI
```

```
EVHLRGRGKLSVLLIPXGEAYCVVTAATPQVHAAALTIARGDRPRVVGELLYRPGEG-
CYLAHAALCCA-QKRTFEEDFVGVYPTKFAKRLTEKLGPSALXHPVGGROVSBS
LFHCDAVASAPSPYSJLPRFEGVEEBAPEITSS:KHKAIESVVERVS:HKDNLARS
VEKOLDFKOEKSKLSKRSKRSVTFYFGEAVQSGLTRAYFCFNLSTFHSVISHPAR
AGSLJENETUASAKSSFDIGGCPCLPHIKGSDTDHVCRIPIYDMKDAQRASRELCQ
ARGLVENIREQVLEAQRVSCHPTLKNVSKDVLIVGVQYDAINEIASAMVJKE
SKVAYLTVPGLLEDEREAFADALGCVVDTRDMQVKFPFGSSCYCHKLNIKMI
MLTPTFTSGNLFSEVMENMGVNYKIKTSAYSPBIRGVKTLRYRACETVVCVKL
PREDKTLFLSGDYIYLDKAFYSRVFDYVUSNCVSVNSKTFEWMWSYIKSKSRVY
ITGKVIHROVDLKHSCFRAVMLAVGVRSTTEFLAKNLNITYGASCFETIRFL
FNEGRVIAEINRDLKMSILSAGUDYBFLDNLNLOHLLEYSEVEVRVSAQNG
VEVCKENRVAJTEIAEADAKRSIAQSGSALSVPTQFGRJGRRGSRVSVSLKYNL
VEEVGNLFTSGVDATFVLVTKFTSDSGIPRVVIMFLDLAAZASPFGVSVSLCAWLR
EAVASFGSWADRTVESVKTFRNTRKFLNFMASLTLLTKKFRFFSASALAKTUV
RKAKVILAEYNEVWPESLSDSGEYSAVEFGSSVITLLTNSGRLLPGFSFSAVITV
LLDLATKISIEVLLKQISFVUSTASSALRYRVLSEILLSNFRMTGHEGIFTKVPFLCCG
LPVVRKVALCVPQDMATYARFLEYGVDDIFFLGRSVNSIKNYLCVVAAGVDSYVD
SVVJLKGAVKVLGFKSGIKIKNPLNFRKAKVTRTSSSTOLSEDEYFSCDSKPG
LRGSSRFTLGRLLDIFENLKSILVIAENACFSAYERIERNMKLYFPPLNSBEEAR
RLRICAGDFDYSDSAPDEMLROAFQYVSSDDSVTYDGKPTVLRSYLVNSRRFL
ETFCNGPKFVKSVNYFKALYSRLLRVLVWDRNLSDSPGLKGGNEKALLAKELKTCV
ITACCVSIOICLRLILCMGTGACGLVRLFYITYSGTRVLRSVUVAVACPLLVRNE
LDGSDGNTNGSVFRFLVALRKALSAYSALRRKIFEFIGNTHHPEDVAVIE
NEVAPPELPEVDIDVDCFGSDSESVDVASIPRPLRGGRSRRSNFTLSLVKVY
FKLAGRIPLRLFRLENFAYVERLAKSKRTFIGLARLDFNLSLTVVYLLOEYDS
VNAFIDVELVLNLSGVNVLPLVSWVRGSLTKLAEVIVGSGFASFLGRMCCRVSQWC
SSSNAGCNFPVPRVTKGFPVPPSSGSTASMEVLEAESDIRHEVLMSTCRVGSDE
EERPEKTEPGIEHTSEVDPVLRSHSOPLSGGECSEYEDREENBRANLLPHVSKIVSE
PRGLETABRNKRTLGVSEELNAINTSNEQRPPIIVDHSPPSRALTSREFPYLQBL
ALFELSCKLRYIDQLKVANFRNGCCLCOEDFVLKAGGVSGRNSRULPHFKG
HEFCFRSGGLVPYDTSVDTFIHTQTFNVSANLLSGYLSYRTFTTNLSANVJXE
APPGGKTLIKVFCETFKSVNLSILTANKSSREEILAKVNRIVLDEGDTPLQTRR
ILITDSVILNMRGATCKVLYLDECPMVGAGAAVACIEFTKCDSAILFGDSRQIYDR
NELTDAVSLDNLRFVDDSVGVSVYRCPMDVCAMLSTFVPTKVATNLVNSAGQSM
QVRLETSDVDVEYSEDFRNLHLOSEKXDLLKSEFKSRSSVEKPTVLTVHEAGCEY
KRYLVLRKQBEDDFRSENHITVALSRHVESYFVSSRRDRACAIKAVCAVJAVJ
AIRVPTPSFGESTLIDISWNPSTDRSKKASSAPYEVINSFLESVPFGTTSVDFGDYS
SEMGQVPEAGNANVVIDSAPVNSKTDHPQVRSSIRSQAIKPKFSLGENDLYSES
RNYNFTVGERFSGOEFQOQAMVMLERSFLEKAVKVRSDVIATIEKGVRTWMSKRE
PSQALRSLGDKLOKPLLEEELITFTKLMVKRDAKVLDSCLVKKPPHQAQNMIEHKAVN
ALFSPCPDEPKNVITCTNSINIVTWNSTLASIAKEMLGSEHVVNVGEIDSKEDJ
KSQDAFIKSFERTLYSAGFPEDDLDMVQOEYTSNATLLOGLSFVDMQFDSGASN
TWIGNSIVTGLNSIYTNRAFKLVFGSDSLIFSPESPIRNSADACEITGEFTKFL
TPSVYVCKFTMGHTVYFVDPYKLVKLGASKDEVDDEFLFVEFTFRDLTKDL
VDERVIELLTHLSKXYGESQDYAAICAIHC:R3NFSSFPKLYPKVKGWVHYGKL
KFVLRKPFANCFREKFDTAFGEAYFLTYDET"
```

```
9445..9609
/note="small hydrophobic protein; ORF2"
/codon_start=1
/product="p6"
/protein_id="AAF14301.1"
/db_xref="GI:6492369"
/translation="MDCVLRSVLLAFLGFLICLFLFCVVFIWFKYQILFRNTPPSN
EARNRSTVV"
```

```
3609..11405
/note="HSP70 protein homolog; ORF3; required for
cell-to-cell movement"
/codon_start=1
/product="HSP70h"
/protein_id="AAF14302.1"
/db_xref="GI:6492370"
```

```
/translation="MVVFGDLFGTTFSSVCAYVGEELYLFKQDASVYPTFVFLHSPT
QEVAFGYDAEVLNDSPVVRGDFYDLKRWIGDCEYNGDYLEK:KPKYKTELLKVASQ
SKSTVRLDCYSOTVPQNAITPLGLATFVKALISTASEAFKQCQCTGVCSVPANVC:Q
RSFTESCYNLGSYPCVYMWNPESAALSACSRIGKATSPVLAVDFGCTGVFVISAQ
NNTFVVRASGGDMGLGSDIDKAFVHEHLYKKAQLPNYKIDISFLKSLSKSVFLNF
PVVSEQNVKDVVLNVSELAEEVAPVETIKIVKEVTEKRGSRMRLEPSPVAKLLKY
GSSVYLPGLLRSLSSVPEBCLVLPDARAAGVAGCALYSACLRNDSFMLVADCAAHN
LSISSKYCEISVCVPAGSIPFTGVTRVNMAGSNASVSAALFEGGFVCKRLNRF
SGDVALGVGVGFSNTRVPLTLEINVSSTVGTITFSLVGPVGVKKLVGNGNAVDFSSY
QLGERVWADLHKHNSDKVKLIHALTYKPEQRKCLTDSKALFLKRLSLADYRERAGKFS
SYDDAVTNSSELGRLRVPKILRGSRVKEVDV"
```

11311..12972

```
/note="64 kDa protein; ORF4"
/codon_start=1
/product="p64"
/protein_id="AAF14303.1"
/db_xref="GI:6492371"
/translation="MTTRSTPANYNYGELFRFFFGGWRKNLMSFAASVRRPVPYDRQ
FNVSENVVDRKNSFSDSTGESFVRFSFLLLPKTYEVCKL:GVAMEQALSQMNRLSD
FVNSLNVVDTKVTGCKENIQTUTVRFKLNGN:AEPSLVEHMSLSNSCGELIKPKD
TRFVSLIFPKDQVVESTDEAVVSSYLDYLSHCLNLYETCNLSSNGKALYDFELK
YVIVLETSDLETRSUSDNPLVAGVLYDMCFYNTJSTYLNKNIESFCFSLYDPLJ
SEIFSMWQOPADPVALLFELDTTELLXIPINTHDSTFLYKNKIRY:LESYFEDDSN
EITKVKVDSLLTRDNPELKAQWVGFCYVGFRTAQTRKVKRDAEYKLPALGEFT
INMSGVEEFDLCKQKQPSVSVRRPCGSLSHEAFSIFKRPFGVGPPI:TRLNVPVKYS
YLVNDYVRYKVRAGLTQDELTILSNLEPDAEMCCEREVALOARQORKEKFPQGWKG
VKNESVSIRVKSNSLNLVNLKVDVGARRQGRNLNPLRKH"
```

```
12920..13570
/notes="ORF5; CPM"
/codon_start=1
/product="minor capsid protein"
/protein_id="AAF14304.1"
/db_xref="GI:6492372"
/translation="MLAPEGRGLIHPTENTRDAMETFNSYDYLAEYSEVNNKLNK
ETDELGVIRERFKSELVITDEDFVKHLAFALIRANITTSKVNYSVAYEYTIQGGK
FLVKDAWVPLJ:KECKKFNKPNVRTFCATFEDAYIVIARSLPKLFLNRTIGKRGIP
SGYFEGADPL:ATSVCLNDHEKAIVLQASRAAIDRAVSSVSDGKIVSLFDLGRLS"
```

```
13641..14255
/notes="ORF6; Cp"
/codon_start=1
/product="major capsid protein"
/protein_id="AAF14305.1"
/db_xref="GI:6492373"
/translation="MGSABPIGAITAFENVSLADQTC:HLHGECDCFLRRDFECLKLG
VPEDKGLAGLGLCLYSKATIGT:SNKVSVOPTSTPIKASFGSGKELFTHGELRSFLDS
QKLLEGKPNKLRCPQRTFOKDYISFAKEYRGLPLPIARANRRLPAEDHY:LAADPST
STELTDOQGRLLLRANHTFESSBPVTSUKQLGRGLATGK"
```

```
14252..14785
/notes="20 kDa protein; ORF7"
/codon_start=1
/product="p20"
/protein_id="AAF14306.1"
/db_xref="GI:6492374"
/translation="MTSSVELLAQTRPLFRVLLKGFVYIVAIETEESEPEVELPLVY
LHDLNENAKKGIKESSYIDFSCMTRLXPSVSVTRVSSGXPSDFSSSYSGKTFDS
KLNLRKVTTFESGIIQVFGMYGRDQRCVSBYLWPFENVFVGAHCGCTLPYCLNCELDK
SGCEILFITSKNEVL"
```

```
14769..15302
/function="enhancer of RNA accumulation"
/notes="21 kDa protein; ORF8"
/codon_start=1
/product="p21"
/protein_id="AAF14307.1"
/db_xref="GI:6492375"
/translation="MKKEFENDCETSRASRSSEL:PRVXELGCTNSPOSSEVSEKINEFN
ELARFNHLVTVHEHMERMEKHQKQSELRTPRJGEMLKEIRAFKRVTVVPMHKEA
SETLNAFLEECYRITGLTREDALREKVRKRVSTVLFHSELKLFVETNMFSTELLK
LNLRLVSISS:LGIAV"
```

```
BASE COUNT 3866 a 347: c 3697 g 4434 t
ORIGIN
```

```
Alignment Scores:
Pred. No.: 6,05e-21 Length: 15468
Score: 297.00 Matches: 71
Percent Similarity: 54.50% Conservative: 32
Best Local Similarity: 37.57% Mismatches: 84
Query Match: 29.97% Indels: 2
D5: 14 Gaps: 2
```

US-09-613-486-15 (1-198) x AF190581 (1-15468)

Qy 7 Ser:AsnLeuSer:Asn:euVallieThr:AspAlaSer:SerLeu:AsmGlyVal:AspLysLys 26

Db 13671 GCGACTTTTGAACACGCTGAGTCTTCGACGACCAACGCTGTTTCACCGTGAAGACTCGGAC 13730

CDS

```

QY 27 LeuLeuSerAlaGluValGluLysMetLeuValGluLysGlyAlaProAsnGluGlyIle 46
DB 13731 AAATCGGAGGATTCGAAGAGTGTGAAATGAAGGGGTTCCGGAAGATAAATC 13790
QY 47 GluValValPheGlyLeuLeuLeuValAlaAlaAlaArgThrSerProLysVal 66
DB 13791 GGCTCGCGGTAGGACTTTGTGTTACTCTGTCGCGACGATAGTACTTCTAATAAGTT 13850
QY 67 GluArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---ArgAsnValVal 85
DB 13851 AGTGTCACACCGAGCTCTACTTTCATCAAAAGCTTCGTTCGTAGTGGGAAGGATTTGTTTC 13910
QY 86 ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105
DB 13911 CTCACTACAGGTGAAGTCTCTTCTGAGTCTTCAGAACTTTTAGAGGGAAGGCT 13970
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaValValAspPheCysIleAla 125
DB 13971 AACAAATTGGGTGTTCTCGCGCACTTTTCAGAAAGGATTACATATCTTTCGCGAAGGAA 14030
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
DB 14031 TACCGAAGAGAGTCCGATTCGTAGAGCAACCGTCACGGTCTACTCTGCTGAAGAT 14090
QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSer 165
DB 14091 CACTACTTACTGCTGATTCATATCGACATCACAGAACTTACTGACCTACACAGGT 14150
QY 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185
DB 14151 CGTCTGCTGTGGCGCGGAAACCGCACACACAGAAATCTCG---TCTGAATCACCA 14207
QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194
DB 14208 GTAACAGTTTGAACAGCTGGGTCGT 14234

RESULT 7
BYU71295
LOCUS
DEFINITION
Beet yellows virus coat protein gene, partial cds.
ACCESSION
U71295
VERSION
U71295.1
KEYWORDS
GI:1619939
SOURCE
Beet yellows virus
ORGANISM
Beet yellows virus
Virus; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
REFERENCE
1 (bases 1 to 612)
Cremer,R., He,X.H., Yang,C.H. and Grantham,G.
Characterization of the 3'-proximal encoded proteins of beet
yellows closterovirus
Unpublished
REFERENCE
2 (bases 1 to 612)
Cremer,R., He,X.H., Yang,C.H. and Grantham,G.
Direct Submission
Submitted (19-SEP-1996) Plant Pathology, University of California,
Riverside, CA 92521, USA
JOURNAL
Location/Qualifiers
FEATURES
source
1..612
/organism="Beet yellows virus"
/mol_type="genomic RNA"
/isolate="California BYV-CA"
/db_xref="taxon:12161"
1..>612
/codon_start=1
/product="coat protein"
/protein_id="AAB17001.1"
/db_xref="GI:1619940"
/translation="MGAEPISAIATENVSLVDOTCLHGDCDCKLRNFEELKLG
VPEDKGLA-GCLYSCTATGTSNKSVQPTSTFFKASFGSGKELFTHGEURSFLOS
QKLEGPKNLRCFRFTQKDYISFAKEYRGRLPPIARANRHGLPAEDHYLAADFIST
STELTDQOGRLLARENATHTEFSSDSPVTSIKQLGRGLATGD"
BASE COUNT 165 a 144 c 147 g 156 t

```

```

ORIGIN
Alignment Scores:
Pred. No.: 1.09e-22 Length: 612
Score: 296.00 Matches: 71
Percent Similarity: 54.50% Conservative: 32
Best Local Similarity: 37.57% Mismatches: 84
Query Match: 29.87% Indels: 2
DB: 14 Gaps: 2

US-09-613-486-15 (1-198) x BYU71295 (1-612)

QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerSerLeuAsnGlyValAspLysLys 26
DB 31 CGGACTTTTGAACCGTGAGTCTGTAGACCAACAGTGTTCGACGGTGAAGACTCGGAC 90
QY 27 LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46
DB 91 AAACCTACGAGGAAATTTTCAAGAGTGTTTGAATTTGAAGGGGTTCGGAAGAATAAATC 150
QY 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLysVal 66
DB 151 GGTCTCGCGTTAGGACTTGTCTGTACTCTCTGTGCGACGATAGGTACTTCTAATAAGTT 210
QY 67 GluArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---ArgAsnValVal 85
DB 211 ACGTCCAAACCGAGCTCTACTTTCATCAAGCTTCGTTCCGCTAGTGGGAAGGATTTGTTTC 270
QY 96 ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105
DB 271 CTCACCTACGCGCAACTGAGGTCCCTTTCTGCACTCTCAGAAACTTTTAGAGGGAAGGCT 330
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAla 125
DB 331 AACAAATTGGCGTGTTCGCGCACATTTTCAGAGGATTACATATCTTCCGGAAGGAA 390
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
DB 391 TACCGAAGAGAGTCCCTCCGATTCGTAGAGCAACCGTCACGGTCTACTCTGCTGAAGAT 450
QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSer 165
DB 451 CACTACTTACTGCTGATTCATATCGACATCACAGAACTTACCGACCTACACAAGGT 510
QY 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185
DB 511 CGTCTGCTGTGGCGCGGAAACCGCACTCACAGAAATTTCTCG---TCTGATTACCA 567
QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194
DB 568 GTAACAGTTTGAACAGCTGGGTCGT 594

RESULT 8
AF056575
LOCUS
DEFINITION
Beet yellows virus Californian isolate, complete genome.
ACCESSION
AF056575
VERSION
AF056575.1
KEYWORDS
GI:3283076
SOURCE
Beet yellows virus
ORGANISM
Beet yellows virus
Virus; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
REFERENCE
1 (bases 1 to 15468)
Peremyslov,V.V., Hagiwara,Y. and Dolja,V.V.
Genes required for replication of the 15.5-kilobase RNA genome of a
plant closterovirus
J. Virol. 72 (7), 5870-5876 (1998)
JOURNAL
MEDLINE
98285746
FJMED
9621048
REFERENCE
2 (bases 1 to 15468)
Dolja,V., Peremyslov,V. and Hagiwara,Y.
Direct Submission

```

```
JOURNAL Submitted (31-MAR-1998) Botany and Plant Pathology, Oregon State
University, 2082 Cordley Hall, Corvallis, OR 97331, USA
FEATURES
Location/Qualifiers
1..115468
/organism="Beet yellows virus"
/mol_type="genomic RNA"
/isolate="Californian"
/db_xref="taxon:12161"
-08..9381
Gene
/gene="ORF1a/b"
CDS
join(108..7983,8045..9381);
/gene="ORF1a/b"
/codon_start=1
/product="fusion protein"
/protein_id="AAC25115.1"
/db_xref="GI:3283077"
/translation="VAFILNVSAVSCAPAPAPAPAHAGASPIV2SPFCVPRYSIDISH
FRILUSLDSVPRPLFNARVHLRASTNPUSPLPLGHASTFVLELNGSSAFPSIPS
RHIDFVNRPSVFTVEVLSSTLPSRLFALLCLDFYCGKEGPCVEIASFSTPP
CLVSNCAQIPTHAMESIRFETKLPAGRLQLQHKRYTKRPETLIIHESGLAKTS
ALGVTSKPSRPI TVKSASGKYEAYE:SRKDFERSRRQOTPRVRSHKPKINAVK
PFFPEPKDKRKRASLPTKDEGPIFTGLRFLPSETPKSEPRLPKPREVEIPVVK
HAPRAVSKPVTPTRPVATTGAEVYNAPTQCSRPRHPILRSASYTFGPKMPLQRP
MKEKDYVYKRSKVSSVTSKPLEALTSILKNLPRTYSERLKFYDHF:GDQFEI
SVPLRGKGLVILLIPKGEAYCVVTAATPQHAALTIIARGDRPRVGGELQYRPGEL
CYLHAALCCALQKRTFREDFVGMYPFTKFAKRLTEKLGPSALKHPVRGRQVRS
LPHCDVASAFSPFYSLPRFIFGVBEAEPIITSLKHAIESVYVRSIHPNDLLARS
VEKD:IDFDEIKSLSKKRSVTVPMYGEAVOSGLTRAYPQFNLSFTHSVYSDHPAA
AGRIILENETIASMAKSFSFDIGGCPFLHIKRGSTDVHVRP:VYMKAQAPRVSRELO
ARGLVEN:SRQELVEAQRVSCVCHTLGNQVKSQVLLIMQVYDASLNEIASANWLEQ
SKVALVNTPGSLDEAREAFDALGDCDVVDTRQVQYFGSGCSYCHLKSINKI
MLTPAFTSGNLFVSEMYENRGVNYKITRSAYSPEIRGVKTLRYRAACTEYVQVNI
PRFDLTKTFLSGYDIYLDKAFVSRVFDYVYVNSGVNSKTFFEWASY:KSSRSIRPL
:SGKVTHRDVHIDKHSECFACAMLAAGVRSRGTTFEFLAKNLYVTSVGAECFTRPL
PREMSRAYAEINRRGFKLMKSLISAGLQDEFLDNLNQLHLEYSEYEVRSIAQNG
EVDNENRVLTEITAEADRKSIAQGLSGLSNVPCOPRGGLGGSRBSGVSLYVJ
VEEVGNLFFSGVDARFLVKVFKTSDSPIFRVRMLD:JAAASPFVSVSLCAWR
EAVSFSWADRTVSBSVKTVPNRTVKRFLNFMASKTLTKKFFFLSALAKATVY
RAKVL:LEAYWEVSESLSDSGESYAVEFCSSVITLTLNSGALLPGFSPSATEY
LLDLATKISIEYLLKQISPVSTASSALYRLVLSINFRMTGHEGIFTKVFV:LCGF
LPVFRKVCALCPGDMATYARFLEYGVDD:FLGSRVNSIKNYLCVVAAGLVDSIVD
SVVLKSGVAKERVLFGSK:IKNF:NVPRKAKVTVTRSSSTDISEDEYFSCDESKPC
LRGSSRRTLSLELLDFENELKSSKLVENACFSAYERIERMKLIYFPPLNSSEBAR
RLIRACGDFDYSDSAFDEDEMLRQAFQYSSDDESVTDGKPTVLRSLYLNSSERRJ
ETFCVCSQICCLRLIRLCWTPACGLVRJFYITYTSGTRV:SRVYVAVAVCLP:VRNE
LDG:SDPLTNGVSRFLFVALRALSAGNSLRKIFEFIFGNLHPDVAIVET
NEVAPGLSPED:IDVDCDQSDSSESVSSDEVASIPRFLGGGSRSENPLTS:VKVY
FK:ACRI:RLL:RLLPNFVAYFVERE:ASKRLKTRIGLAF:FNRLSLTSVWVLLQEYS
VLNAT:IDVELV:JNSGSVNVLPVSVKRGSLTKAEVIVGSGFASFLGRXC:CRVSDWC
SSSRAGCNFMSPVRTKGFVPPSSGSTASKYERLEADESDIREHVLSTCRVSDDE
EERKPEVTEPGIEHTSEVPIRSHSQPSGEGCSYSEDREERANLLPHVSK:VSE
RGLTE:ARNRKT:LRGSEFLNAIN:SNQRPRIIVDSHSPESRALTSVRFFYLQEL
ALFELSKRLREYDOLKVANTRQBCLOCDKDEPVLGAGQVSGVSRNRLPLKHXID
HEFCFRSGGLVPYDGTSDVTIIFHTQTFVSNAL:LSGY:SVYRFTFTN:SANVLLYE
APPGGKKTTLIKVCFEFSKVNSLILTANKSSREELAKNRIVLEGGDTPJQTROR
LITUDIS:LNMRKL:CKVLYIDECFVSVHAGAAC:EFTKCDSAILFGDSQIHYIDR
NELO:TAVIDLNRFDVDESVYGYRCPMDVCWALTTPYKTVATIN:VSAQSSM
QVREISVDVDEYSEFVLTMLQEKDRLSKFGKRSRSEVKEPVLTVHAEQGYT
RKVNLRVTKFOBDQCP:RSENH:TVALSRLHVESLTSVLSLSKRDDAIAQAI:VKAQLVD
AVRYVPTSGGSTLDISVNPSTSDRSCKKASAPVEVINSFLESVWPGT:SVDFGQVS
EEMGTQVPSGADNVIRDSAPVNSTDHEENLYSIESRNNFTVCRFPQGFQGA
MANW:LEKRS:DLERKAVKRSQVIAITEKGRTWMSKREPSQRLASLSDQLKPLNEE
IITFFLMTNNTSLAGIAKEMLGSHEVYVNGEIDFSKDKSQDAFKSPERFLYASGFG
DEDDLWVWQGYSTNATLDQLSFSVDNQRKSGASNTWIGNSIVTLGILSMFYVTN
RFKALFVSGDSDSLTSESPISNADAXCTELGFEKFLTPSVYFCSKFFWMTGHPF
FVPDPYKLLVKLGASKQSDVDEFLFEVFTISFDLLTKOLVDIERVIELLTHLVSHKGYE
SGDTTAAALCAIHCIRSNFSFKKUKYPKVGMWVYHVKLFVLRFKAFNCFREKFTAFG
EAYFLTYDET"
108..1871
/mat_peptide
/gene="ORF1a/b"
/product="papain-like leader peptide"
```

QKLEKPKNLCRCFCRTFKDYISFAKEYRGRLPPIARANRHHGLPAEDHYLAADFIST  
STELTDLQCGRLALARENAHTTFSSSPVTSKQLGRGLATGK"  
14252..14785  
/function="unknown"  
/notes="ORF 7"  
/codon\_start=1  
/product="20 kDa protein"  
/protein\_id="AAC25121.1"  
/db\_xref="GI:3283083"  
/translation="MSSVELAQTKPLFRVLLKGFVFIVALETBEESPEVELPLVY  
LHDLNLNKKGISYIDFKSMTRLKPSVSYRVSSGKPSSEDFSSYSGKTFDS  
KILNKVTFESGICLGFQMGQRDCRSVSEYLFNFVFGAHCGTJYCLNCLCK  
SGSELELTFSEKNEFL"  
14769..15302  
/functions="replication enhancer"  
/notes="ORF 8"  
/codon\_start=1  
/product="21 kDa protein"  
/protein\_id="AAC25122.1"  
/db\_xref="GI:3283084"  
/translation="MKFFPNDCTSRASRSSELLRRVKELGTNSPQSEVSECINEFN  
ELARPNHLVTVVHRERMEKHPKQSELRLPSRLGEMLKEIRAFJAVRVTPMHEKTA  
SETLNAFLSEYCYRTGLTRDALREMRKRVSTVLFHSELUXFEVTENKFSFTELLK  
LNLSJRVISSQILGIAY"

CDS

BASE COUNT 3874 a 3472 c 3696 g 4426 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9,93e-21 Length: 15468  
Score: 295.00 Matches: 71  
Percent Similarity: 53.97% Conservative: 31  
Best Local Similarity: 37.57% Mismatches: 85  
Query Match: 29.77% Indels: 2  
DB: 14 Gaps: 2

US-09-613-486-15 (1-198) x AF056575 (1-15468)

Qy 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26  
Db 13671 GCGACTTTTGAAACGCGAGTCTCGTAGACCAACAGGTGTTTCACGCGTAGACTCGGAC 13730  
Qy 27 LeuLeuSerAlaGluValGluLysMetJeuValGlnLysGlyAlaProAsnGluGlyIle 46  
Db 13731 AAATCCGAGGAAATTCGAAGAGCTGTTGAAATGGAAGGGTTCGGAAGATAAATC 13790  
Qy 47 GluValValPheGlyLeuLeuLeuTyTAlaLeuAlaAlaargThrThrSerProLysVal 66  
Db 13792 GGTCCTCGGTAGGACTTGTGTTGACTCTCTGTCGACGATAGTACTTCTAATAAGTT 13850  
Qy 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---ArgAsnValVal 85  
Db 13851 AGCGTCCACCGAGCTACTTTCATCAAGCTTCGTTCGCTAGTGGAGGAATGTTTC 13910  
Qy 86 ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105  
Db 13911 CTCACCTCAGCGCAACTGAGGTCTCTGACTCTCAGAACCTTTTAGAGGAAAGCCCT 13970  
Qy 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyTValAspPheCysIleAla 125  
Db 13971 AACAAATTCGGTGTCTGCGCGACCTTTTCAGAGGATTAATATCTCTCGCGAAGGAA 14030  
Qy 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145  
Db 14031 TACCGAGGAAAGCTCCCTCCGATTCAGAGCAACCGCTCAGCGTCTACCTGCTGAGAT 14090  
Qy 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165  
Db 14091 CACTACTTAGCTGTGATTTTCATATCGACATCAACAGACTTACTGACCTACACACAGGT 14150  
Qy 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185  
Db 14151 GGTCTGCTGTGGCGCGGAAACCGCACTCACAGAAATTCGCG---TCTGAATCACCA 14207

Qy 186 ValSerAsnLeuArgGlnLeuGlyArg 194  
Db 14208 GTAACTAGTTGAAACAGCTGGTCTG 14234

RESULT 9

BYVCPG 2724 bp ss-RNA linear VRL 02-AUG-1993  
Beet yellows virus coat protein gene, complete cds, and two ORFs.  
DEFINITION  
ACCESSION M59452  
VERSION M59452.1 GI:323237  
KEYWORDS coat protein.  
SOURCE Beet yellows virus  
ORGANISM Beet yellows virus  
Viruses; ssRNA positive-strand viruses, no DNA stage;  
Closteroviridae; Closterovirus.  
REFERENCE 1 (bases 1 to 2724)  
AUTHORS Brunstedt,J., Moseley,J. and Hull,R.  
TITLE Nucleotide sequence of cDNA encoding the coat protein of beet yellows virus  
JOURNAL Unpublished (1990)  
COMMENT Original source text: Beet yellows virus cDNA to genomic RNA.  
FEATURES  
source  
1..2724  
/organism="Beet yellows virus"  
/mol\_type="genomic RNA"  
/isolate="Broom's barn"  
/db\_xref="taxon:12161"  
CDS  
1..1287  
/codon\_start=1  
/label=ORF\_1  
/protein\_id="AAA72953.1"  
/db\_xref="GI:323238"  
/translation="PDFVRLKLNLAEPSSLVEHCHMSNSCGELINPKDKRFPVSLIF  
KGVVESTDEAVVSSVLDYLSHCLNLYETCNSSNGKALYDEPLKYVIAVLETS  
DLEVRSPDNPLVAGLYDMCFEYNTLSTYLNKRLAYSEFDDSNELIKVKVDS  
QPADVRLLELDTAELSLKIPTINTLSTFYLNKRLAYSEFDDSNELIKVKVDS  
LTDNPELKLQARWVGPHCYGVFRFTRQTRKVRDAEYKLPALGETFINMSGVVEF  
FKDLQKPVSVRRFCGSLSEAFSIFKRGVGFPPITRLNVPVKYSYLNVDYRHH  
VRAGLTODELTILSNIBFVAEMCCEREVALQARRQGEKPFQGMKGVKNEVSPHA  
RSSIRVAKNSLNLILWKDVGARRQRLNPLRRKH"  
1235..1885  
/codon\_start=1  
/label=ORF\_2  
/protein\_id="AAA72954.1"  
/db\_xref="GI:323239"  
/translation="WLAPEGRDLTHFTENTRDAMETFNSSVDLAEYSEVNPKNLNRK  
ETDELGVIRPRKSELVITDEDFVKHLAFALIRAGN:TTTKNVYVAYEYTIQGX  
FLVDKAWFPLIKCKMKFKNPNVPTFCATFEDATVIARSLPKLNLRNRTIGKGP  
SGVEFLGADFLUTATSVCLNDHEKALVLOASRAAIDRAVSSVDGKIVSLFDLGRUS"  
1956..2570  
/codon\_start=1  
/label=ORF\_3  
/product="coat protein"  
/protein\_id="AAA72955.1"  
/db\_xref="GI:323240"  
/translation="MGSAPISATATFENVSLADQTLHGDCCKLRKXNFECLKLG  
VPEDKLGLALGLYCATTGTSNKSQVPTSTFIKASFGGKELFTHGLRSLDS  
OKLSEGNKLRKCFQRTFKQDYISFAKEYRGRLPPIARANRHHGLPAEDHYLAADFIST  
STELTDLQCGRLALARENAHTTFSSSPVTSKQLGRGLATG"

source

CDS

CDS

CDS

BASE COUNT 739 a 591 c 638 g 756 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,64e-21 Length: 2724  
Score: 293.00 Matches: 71  
Percent Similarity: 53.97% Conservative: 31  
Best Local Similarity: 37.57% Mismatches: 85  
Query Match: 29.57% Indels: 2  
DB: 14 Gaps: 2

US-09-613-486-15 (1-198) x BYVCPG (1-2724)

Qy 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26

```
Db 1986 CGGACTTTGAAAGCTGAGTCTCGCAGACCAACGCTGTTCGACGCTGAAGACTCGGAC 2045
Qy 27 LeuLeuSerAlaG:uValGluGlyMeLeuValGlnGlyysGlyA:aProAsnGluGlyIle 46
Db 2046 AAATACCGAAGAA:TTTCGAAGAGTGT:TTGAAATGGAAGGGGTTCGGAAGACAAATC 2105
Qy 47 GluValValPheGlyLeuLeuTyTAlaLeuAlaAlaArgThrThrSer?roLyysVal 66
Db 2106 GGTCTCGGCT:AGGACTTGTGTGTATTCCTGTCGACGATAGTACTTCTTAATAAGATT 2165
Qy 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal: 85
Db 2166 AGTGTCCAACCGACGTCTACTTTTCATCAAGCTTCGTTCGCTGGTGGGAAGAAATGTTC 2225
Qy 86 ValThrGluGlyAspLeuLySlysVal:LeuAspGlyCysAlaProLeuThrArgPheThr 105
Db 2226 CTCACCTACGGTGAACCTCAGGTCTTTCTTGACCTCTCAGAACTTTTGGGAAGAAAGCCT 2285
Qy 106 AsnLySlysLeuArgThrPheGlyArgThrPheThrGluAlaTyValAspPheCysI:eAla 125
Db 2286 AACAAAGTTCGCTGTTCTGCGGCACITTTTCAGAAGGACTACATATCTTCGCGAAGGAA 2345
Qy 126 TyLyShisLySlysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 2346 TACCGAAGAAAGACTGCCTCCGATTGCTAGAGCCCAACCGTCACGCTCTACCTGTGAAGAT 2405
Qy 146 SerTyLeuAlaAlaAspPheLeuGlyThrCysProLySlysLeuSerGluLeuGlnGlnSer 165
Db 2406 CACTACTTAGCTGCTGATTCTCATATCGACATCACAGAACTTACTGACCTACCAAGGT 2465
Qy 166 ArgLySlysPheAlaSerMetTy:AlaLeuLySlysThrGluGlyGlyValValAsnThrPro 185
Db 2466 CGTCTGCTGTTGCGCGCGGAAAGCCACTCACACAGAGTCTCG---TCTGAATCACCA 2522
Qy 186 ValSerAsnLeuArgGln:LeuGlyArg 194
Db 2523 GTAAGTATTGAAACAGCTGGTGGT 2549

RESULT 10
BYVMBPA
LOCUS
DEFINITION
Beet yellows virus (isolate German BYV-3) genes for putative
membrane-binding protein, heat shock 70-related protein, coat
protein homolog and coat protein.
ACCESSION
X73475.1 GI:313689
VERSION
X73475.1 GI:313689
KEYWORDS
coat protein; heat shock 70-related protein; membrane-binding
protein;
SOURCE
Beet yellows virus
Beet yellows virus
Beet yellows virus
Virus; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
REFERENCE
1. Agratovsky,A.A., Koonin,E.Y., Boyko,V.P., Maiss,E., Lunina,N.A. and
Atabekov,J.G.
Beet yellows closterovirus: complete genome structure and
identification of a leader papain-like thiol protease
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 5980)
TITLE
Direct Submission
JOURNAL
Maiss,E.
Pflanzenvirologie, Biologische Bundesanstalt, Institut fuer
Biochemie, Messweg 11/12 D-3300 Braunschweig, FRG
FEATURES
Location/Qualifiers
1. .5980
/organism="Beet yellows virus"
/mo_type="genomic RNA"
/isolate="German (BYV0G)"
/db_xref="taxon:12161"
/clone="pBYV-D22, pBYV-D31"
71. .235
CDS
```

```
/note="putative membrane-binding protein"
/codon_start=1
/protein_id="CAA51857.1"
/db_xref="GI:313690"
/db_xref="SPTREMBL:Q65893"
/translation="MDCVLRSYLLAFGLICLFLCLVVFIFVYKQI--FRNPPSN
EAREFRSIVV"
235. .2031
/codon_start=1
/product="heat shock 70-related protein"
/protein_id="CAA51858.1"
/db_xref="GI:313691"
/db_xref="SPTREMBL:Q65894"
/translation="WVVG-DRTFTSSVCAYVGEELYFKORDSAVIFTFVLSHSDT
QEVAFGYDAVLSNDSSVRGCFYRDILXRNWIGCBENYGDYLEKLKPHYKTELLKVAQS
SKSTVKDCDCTSGVTFQNALTPGLIATFVKALISTASEAFKQCTGTGVCISVFNANCQ
NFTFESCVNLGSPCYVYVMEPEPSAAASACSRKIGATSPVLVYDFGGTTFDVSISAL
RSFTVRAASGDMNLGGRDIDKAPFVHLRYKQLPVNYKIDISFLKESLSKKVSYFLNF
PVSEQNKYVDVLVNVSELAEVAAPFVERTIKIVNEVYKRGSMLEPESVKACLNV
GGSYLQGLLSRLSSVPFVEECVLDPARAAGGACALYACLRNDSPMLLVCAAHN
LS:SSKYCESTVCVPAGSPIPFTGRTVNMTCGNSASAVYSAALFEGDFVKRLNKRIP
SGDVALGDVGFGNSNTRTVFLTLEINVSSTGTTTFLVGPTVKKLVGNAADFFSY
QUGSRVADLHKNSDKLIHALTYKPFORKKLTDSKALFLKRLSADYRRSARKEP
SYDDAVLNSSELLLGRVIPKILRGRVEKLDV"
1937. .3598
/note="orf4"
/codon_start=1
/protein_id="CAA51859.1"
/db_xref="GI:313692"
/db_xref="SPTREMBL:Q65895"
/translation="MTTFTSTPANYMGELFRFFGGREWNLMSSEASVRRFVYARD
FRFSGVLRKIPFSDTGESFVREFSLILTPFKTEVCKLCGVAMEQALSGMNRLSD
YNSVLSNVAVDKTVGCKFN:QTVTEFVKLGNLALPS:VEHCWSLSNSCGELINPKD
TKRSLIIFKGQVIESDTBAIVSSVLDVLSHCLNLAYETCNSNSGKKALYDEFLK
VVIAYLSDLEYHSPSDNLVAGVLYDMCFEYNTLKYTKNIESPCDCLSLYLPJL
SEIFSNWBPAPDRLFLDLTAEALLLKIPTINTLDSTELYKNKRYESIFEDDSN
ELIKVQSLSLTRDNPDLKLAQRWVGPHCYTGVFRTAQRKVRKRDAYEIKPCPALGET
INMSVEEFDELQCKMFSVRRFCGSLSEAFSIFKRFVGVPPIITLNPVFKYS
YLNVDYRHVKRAGLTQDELITLSNIEFDVAEMCCEREVALQARRAQRREKPKQGNKG
VRNEVSPHARSIRVKKSNESLLNLWKDVGVRROGLNPLRKH"
3546. .4196
/codon_start=1
/product="coat protein homolog"
/protein_id="CAA51860.1"
/db_xref="GI:313693"
/db_xref="SPTREMBL:Q65896"
/translation="MLASGRGLIHFTENTRCAMETFFNSYDLAEYSEVNPKNLNRK
STDELLGVRERPKSELVITDEDFVGHAPALIRAAINI:TSKVNYVGAVEY:IGSKK
FLVKDAMVFLIKCKMKKFKNRPVTFCA:TFEDAYIV:ARSIPKFLARTIGKRI:
SGYFLGADRLTATSVCLNDHEKAIVLQASRAAIDRAVSSSDVKIVSLFDLGRLS"
4257. .4881
/codon_start=1
/product="coat protein"
/protein_id="CAA51861.1"
/db_xref="GI:313694"
/db_xref="SPTREMBL:Q65897"
/translation="MGAPE:SAIAIFENVSLADQTCJHGDCDCKLRNFEBCLKJLKG
VPEDLGLALGLCYSCATIGTSNKVSQPTSFIRKASFGGKELF:THGEESFLDS
QKLEGRKNLRGCFRTFKQYISFAKEYGRFLPIARANRHGLPAEDHYLAADFIST
STELTDLQOGRLLIARENA:HTBFSSSPVTSLKQJGRGLATOR"
4878. .5411
/note="orf7"
/codon_start=1
/protein_id="CAA51862.1"
/db_xref="GI:313695"
/db_xref="SPTREMBL:Q65897"
/translation="MIGSVELAQTTRPFRV:LLKGFVFIYVIALETEEPESEAEPLPVY
LHDLNLVNNKGI:ESSYIDFKSCMRLLKPSYSYTRVSGKPSSEFSSYSKTFES
RILNRKVITPENGTLQVPGMYGRDRRCVSSSEYLMWFENVFVGAHCGTLPYCLNCELDK
SGGDLEILTFSKNEVFL"
5395. .5928
/note="orf8"
/codon_start=1
```

viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae; Closterovirus; SYDDAVLNSFFLLGGLPIIKHIGSVGEKVV; OLGENVAVLHKNSDDKVI HANLYPOPEKKLLIDGDKALFVKLL;ADIRREARKKS



CDS

```

2576..4237
/contig="OR34"
/codon_start=1
/protein_id="CAA37552.1"
/db_xref="GI:58881"
/translation="MTTRESTPANYWYMGELFRFRFGQEWKMLMSEAAVSRRPVSSD
FRSDGVILSRKTFGSGTGVSRFESLTLTPKTYEVCKLQVAMEALANGMRLSD
VNVSEFIVDVKTGCKENIQSVTFKINGVNAEPLSVHCHMSLSKCELINPKD
TRPVSLIPKGLDASTDEAVSSYLQYLSHCLNHYETCNSSNSGRKSLIDFJK
HVIDLENSDLRSPSDNPLVAGILYDMCFEYNTLSTYLNIESFCDFSLYLPL
SEFSKMWERPAPDVRLLFELDAEALLKVPFTINMHDQTLKYKRLRYLESYFEDSN
ELIKVVDLSLLTRDNDELKALQWGFHYGVYVTRQTKVKRDAEIKLPALGEFY
INMGVEEPEELCKXKMPISVRRKSGSLSEAFVFKFGVGPPTIRLNVPVKYS
YLVNDYVRKVRGLTQDELTLISNIEFPAVMCCEREVALQARQRERPFQCKG
TKNEISPHARSSIRVKNDLSLLNLIKVDVGARSQRRLNPLRKH"
4185..4835
/contig="ORF5"
/notes="ORF5"
/codon_start=1
/protein_id="CAA37553.1"
/db_xref="GI:58882"
/db_xref="SPTREMBL:Q08538"
/translation="MAPARGDLIHFTENTRDMETPFNSVDLLEYSEVRPNKLNK
ETDELGVIRERKSELVITDEDFVHLAPALIRANITTSVKVNYGVAYETIGGK
FLVKDAVPLIKCKMKXKNPNVTRFCATTFDAYIVTARSPLKFLNRTIGRGIP
SGVEFAGAPLTAISVCLNDHEKAIIVLOASRAIIRAVSSVDGKIVSLFDLGRLS"
4906..5520
/codon_start=1
/product="capsid protein"
/protein_id="CAA37554.1"
/db_xref="GI:58883"
/db_xref="SPTREMBL:Q08531"
/translation="MGSASPIGAIATFENVNSLADOTCLHGDCCKLRKNFBCMLKG
VPEDNLGIALGLCYCATIGTSKNVNOPTSTFIKASFGGCKELYLTHGLNSFLGS
KXLECKPKNLCFCGTROKDIISLRKYGKLPPTARNRHGPREOHYLAADIST
STSLDQSRJLLARENA:HTFEFSSESPTSLKQLGRGJGTGR"
5517..6059
/contig="ORF7"
/notes="ORF7"
/codon_start=1
/protein_id="CAA37555.1"
/db_xref="GI:58884"
/db_xref="SPTREMBL:Q08544"
/translation="MTSSVVELAQTKPLFRVLLKGFVTVVAFTEESSEAEPLVY
LHDFELNKRKGIKSIKSYVFMSCVTRJRPSSVSVTRVSESSDFSLPGSGKTFGS
KVLNRKVTTFENGVLQVFGMYGLEQRCVSSDYWFENFVFGAHGCTLYCINCEJDK
SGGELEILTPSKNEVLLKRW"
6034..6567
/contig="ORF8"
/notes="ORF8"
/codon_start=1
/protein_id="CAA37556.1"
/db_xref="GI:58885"
/db_xref="SPTREMBL:Q08545"
/translation="YKFFLKDGETSALSRSESLLRKVRKELGTNSQOOSRISCVDEFN
ELASFNEJLVTVHREHMQHPNOSSKLRVPSRIGMKLEIRAFKRVVTVWKECA
SDTLNAPLEYCHRTGLAREDALREKRVKSVLPHHSSELKFEVTEWMPSTYELDK
LNLISRVISQILGMAI"

```

CDS

```

BASE COUNT 1775 a 1415 c 1581 g 1975 t
ORIGIN
Alignment Scores:
Pred. No.: 3 08e-20 Length: 6746
Score: 286.00 Matches: 71
Percent Similarity: 52.91% Conservative: 29
Best Local Similarity: 37.57% Mismatches: 87
Query Match: 28.86% Indels: 2
DB: 14 Gaps: 2

```

US-09-613-486-15 (1-198) x CLBVU3PH (1-6746)

Alignment Scores:

```

Pred. No.: 3 08e-20 Length: 6746
Score: 286.00 Matches: 71
Percent Similarity: 52.91% Conservative: 29
Best Local Similarity: 37.57% Mismatches: 87
Query Match: 28.86% Indels: 2
DB: 14 Gaps: 2

```

```

Qy 7 SerAsnLeuSerAsnLeuVali.eThrAspAlaSer:SerLeuAsnGlyValAspLysLys 26
Db 4936 GCGACTTTTCAAACGTAAGTCTCGCAGACCAACAACTGTTGCGACGGAAGACTCGCAT 4995

```

```

Qy 27 LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyLe 46
Db 4996 AAACCTTAGGAGAACTTCGAGAGAGTGTGAAATTAAGAGGGTTCGGAAGATAACCTC 5055
Qy 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAargThrThrSerProLysVal 66
Db 5056 GGAATCGGTAGGACCTTGTGTTGATTCCTGCTACGATAGGCACCTTCCAAACAAAGTT 5115
Qy 67 GlrArgAlaSerAspValIlePheSerAsnSerPhe---glyGluA:ArgValVal 85
Db 5116 AACGTCCAAACCGAGCTACCTTCATCAAAAGCTTCGTTGCTGCTGGAGGAAGACTGTAC 5175
Qy 86 ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr 165
Db 5176 CTCACTCAGCGTGAATTAATTCCTTCTGGGTCTCAAAAACCTTTTGAGGAGGAAACCT 5235
Qy 106 AsnLysLeuAargThrPheGlyA:GThrPheThrGluAlaTyrValAspPheCysIleAla 125
Db 5236 AACCAATTCGGGTGTTTCTGCGGTACTTT-CAGAAAGGACTACATATCTTTCGCGCAAGAA 5295
Qy 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 5296 TACGAGGGGAATTAACCTCCGATTCGAGAGTAACCGTCACCGTCTACCGCTGAAGAT 5355
Qy 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 5356 CACTACTTAGCGGTGACTTCATATCGACGTCGACGGAACCTCACTGACCTACAAACAAAGT 5415
Qy 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185
Db 5416 CGTCTGCTGTAGCGCGCAAAACGCCACTCACACGGAATTCCTCG---TCTGAATCACCG 5472
Qy 186 ValSerAsnLeuArgGlnLeuGlyValG 194
Db 5473 GTACCAAGTTGAAACACACTAGTGTGCT 5499
RESULT 12
LOCUS BYVJAA 15480 bp RNA linear VRL 19-MAY-1995
DEFINITION Beet yellows virus (isolate Ukrainian BYV-U) complete genome.
ACCESSION X73476
VERSION X73476.1 G:405624
KEYWORDS coat protein; heat shock 70-related protein; heat shock 90-related protein; methyltransferase; papain-like protease; RNA dependent RNA polymerase; RNA helicase.
SOURCE Beet yellows virus
ORGANISM Beet yellows virus
Viruses; ssRNA positive-strand viruses, nc DNA stage;
Closteroviridae; Closterovirus.
REFERENCE 1
AUTHORS Agranovsky, A., Koonin, E.V., Boyko, V.P., Maiss, E., Frotscher, R.,
Lutina, N.A. and Atabekov, J.G.
TITLE Beet yellows closterovirus: complete genome structure and
identification of a leader papain-like thiol protease
JOURNAL Virology 198 (1), 311-324 (1994)
NEEDLINE 94082464
PUBMED 8259666
REFERENCE 2 (bases 1 to 15480)
AUTHORS Agranovsky, A.A.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1993) A.A. Agranovsky, A.N. Belozersky Institute,
Moscow State University, 119899 Moscow, Russia
COMMENT Related sequence: X53462.
FEATURES
Location/Qualifiers
1..15480
/organism="Beet yellows virus"
/moltype="genomic RNA"
/isolate="Ukrainian (BYV-U)"
/db_xref="taxon:12161"
/cclone="pBYV p36, x19, R9, R3, 36A, R2, 36, p43, 111-7,
112, 115, 1213, 1210, 124, 1313, 1311, 142, 143, 1311,
1311, 142, 143, 156, 1510, 154, 1518, 1520, 169, 1615, 175"
1..107
5'UTR

```





```

CDS
SGSELEILTSKNEYJLKRW"
14768..15301
/locus="orf8"
/codon_start=1
/protein_id="CAA51870.1"
/db_xref="GI:405633"
/db_xref="SPTREMBL:Q08545"
/translation="MKFFADGTSRSLRSSELLRRVVELGTNSQCSISCVDFN
ELASFNHLLVTVREHMEQKRSKLRVPSRIGEMKEIRAF.LKRVVTVMKETA
SCLNATFEYCRITGLAREDALREKMKVKSVLFHHSLLAFEVTEKMFYSITELLK
LNLRLVISOILGMAI"
15299..15480
3'UTR
BASE COUNT 3891 a 3446 c 3680 g 4463 t
ORIGIN
Alignment Scores:
Pred. No.: 9,23e-20 Length: 15480
Score: 286.00 Matches: 71
Percent Similarity: 52.91% Conservative: 29
Best Local Similarity: 37.57% Mismatches: 87
Query Match: 28.56% Indels: 2
DB: 14 Gaps: 2
US-09-613-486-15 (1-198) x BYVCAA (1-15480)
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26
...
DB 13670 CGGACTTTGAAACGTAAGTCTCGGAGACCAACCTGTTTCACGGAGAGACTCCGAT 13729
QY 27 LeuLeuSerAlaGluValGluMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46
...
DB 13730 AAACCTTAGGAAGAACTTCGAAGAGTGTTCGAAATTAAGGGGTTCGGAGATAACCTC 13789
QY 47 GluValValPheGlyLeuLeuTyTAlaLeuAlaAlaArgThrThrSerProLysVal 66
...
DB 13790 GGAATCGCGTTAGAGACTTTGTTGATTCCTGCTACGATAGCAGCTTCCAAACAAAGTT 13849
QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85
...
DB 13850 AACGTCACACCGAGCTCTACCTTCATCAAGCTTCGTTGCTGGTGGAGGAAGTGTAC 13909
QY 86 ValThrGluGlyAspLeuLysValLysAspGlyCysAlaProLeuThrArgPheThr 105
...
DB 13910 CTCACCTCAGCGTGAAATGAAATCCCTTCCTGGGTCTCAAAACCTTTGGAGGGAAACCT 13969
QY 136 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyValAspPheCysIleAla 125
...
DB 13970 AACAAATTCGGGTGTTCTCGCGTACTTTTCAGAGGAGCTACATATCTTCGCGAAGGAA 14029
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
...
DB 14030 TACCAGGAGAAATTACCTCCGATTCGCCAGAGCTAACCGTACCGGTCTACCCGCTGAAGAT 14089
QY 146 SerTyValLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
...
DB 14090 CACTACTTACCGCGCTGACTTCATATACGAGCTCGAGGAGCTACCTGACCTACCAACAGT 14149
QY 166 ArgLysMetPheAlaSerMetTyTAlaLeuLysThrGluGlyGlyValValAsnThrPro 185
...
DB 14150 CGTCTGCTGTTGAGCGCGGAAACGCCACTCACACGGAATCTCG---TCTGAATCACCG 14206
QY 166 ValSerAsnLeuArgGlnLeuGlyArg 194
...
DB 14207 GTAACAGTTTGAAACAACTAGTCTGT 14233
RESULT 13
A41914
LOCUS A41914 783 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 1 from Patent WO942847.
ACCESSION A41914
VERSION A41914.1 GI:2297469
KEYWORDS Beet yellows virus
SOURCE Beet yellows virus
ORGANISM Beet yellows virus
Viruses; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
REFERENCE 1 (bases 1 to 783)
AUTHORS Bojser,K. and Brunstedt,J.
TITLE DNA SEQUENCE COMPRISING AT LEAST TWO COAT PROTEIN GENES
JOURNAL Patent: WO 9428147-A 1 08-DEC-1994,
SANDOZ AG (AT)
COMMENT OTHER PUBLICATION AU 7122694 941220.
FEATURES
Location/Qualifiers
source
1..783
/organism="Beet yellows virus"
/mol_type="genomic DNA"
/db_xref="taxon:12161"
BASE COUNT 223 a 176 c 171 g 211 t
ORIGIN
Alignment Scores:
Pred. No.: 2.3e-21 Length: 783
Score: 285.00 Matches: 70
Percent Similarity: 53.44% Conservative: 31
Best Local Similarity: 37.04% Mismatches: 86
Query Match: 28.76% Indels: 2
DB: 6 Gaps: 2
US-09-613-486-15 (1-198) x A41914 (1-783)
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26
...
DB 119 CGGACTTTGAAACGTAAGTCTCGGAGACCAACCTGTTTCACGGTGAAGACTCCGAC 178
QY 27 LeuLeuSerAlaGluValGluMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46
...
DB 179 AAACCTTAGGAAGAACTTCGAAGAGTGTTCGAAATTAAGGGGTTCGGAGAAACATC 238
QY 47 GluValValPheGlyLeuLeuTyTAlaLeuAlaAlaArgThrThrSerProLysVal 66
...
DB 239 GGTCTCGCGTTAGAACCTTGTTCCTTCCTGTCGCGACGATAGGTACTTCTTAATAAGTT 298
QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85
...
DB 299 AGTGTCCAACCGAGCTCTACTTTCATCAAGCTTCGTTCGTGGTGGAGGAAGTGTTC 358
QY 86 ValThrGluGlyAspLeuLysValLysAspGlyCysAlaProLeuThrArgPheThr 105
...
DB 359 CTCACCTCAGCGTGAACCTGAGGTCTTTCGACTCTCAGAAACTTTTGGAGGAAAGCCT 418
QY 156 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyValAspPheCysIleAla 125
...
DB 419 AACAAATTCGGGTGTTCTCGCGTACTTTCAGAGGAGCTACATATCTTCGCGAAGGAA 478
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
...
DB 479 TACCAGGAGAACTGCTCCGATTCGTAGACCAACCGTACCGTCTACTCTGCTGAAGAT 538
QY 146 SerTyValLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
...
DB 539 CACTACTTACCGCGCTGATTCATATACGATACACAGAACTTACTGACCTACACAGGT 598
QY 166 ArgLysMetPheAlaSerMetTyTAlaLeuLysThrGluGlyGlyValValAsnThrPro 185
...
DB 599 CGTCTGCTGTTGAGCGCGGAAACGCCACTCACACAGAGTTCG---TCTGAATCACCA 655
QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194
...
DB 656 GTAACAGTTTGAAACAACTAGTCTGT 682
RESULT 14
BYU51931
LOCUS BYU51931 10545 bp RNA linear VRL 05-APR-1999
DEFINITION Beet yellow stunt virus helicase gene, partial cds; and RNA
replicase, p30, p6, p86, p61, p25, coat protein, p18, and p22
genes, complete cds.

```

```

J51931 L20761
US1931.1 GI:1388128
Beet yellow stunt virus
Beet yellow stunt virus
Viruses; serNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
1 (bases 2734 to 5576)
Karasev,A.V., Nikolaeva,O.V., Koonin,E.V., Gumpf,D.J. and
Garney,S.M.
Screening of the closterovirus genome by degenerate primer-mediated
polymerase chain reaction:
J. Gen. Virol. 75 (Pt 6), 1415-1422 (1994)
9267425
9207405
2 (bases 1 to 10545;
Karasev,A.V., Nikolaeva,O.V., Mushagian,A.R., Lee,R.F. and
Dawson,W.O.
Organization of the 3'-terminal half of beet yellow stunt virus
genome and implications for the evolution of closteroviruses
Virology 221 (1), 199-207 (1996)
96266429
8661428
3 (bases 1 to 10545;
Karasev,A.V., Nikolaeva,O.V., Lee,R.F., Wisler,G.C., Duffus,C.E.
and Dawson,W.O.
Characterization of the beet yellow stunt virus coat protein gene
Phytopathology 88 (10), 1040-1045 (1998)
4 (bases 1 to 10545)
Karasev,A.V.
Direct Submission:
Submitted (20-MAR-1996) Alexander V. Karasev, University of
Florida, Citrus Research and Education Center, 700 Experiment
Station Road, Lake Alfred, FL 33850-2299, USA
On Jun 25, 1996 this sequence version replaced gi:507953.
Location/Qualifiers
1. 10545
/organism="Beet yellow stunt virus"
/mol_type="genomic RNA"
/specific_host="Sonchus oleraceus L."
/db_xref="taxon:35290"
<1..2133
/function="replication"
/notes="ORF1a; similar to helicases of other
closteroviruses"
/codon_start=1
/product="helicase"
/protein_id="AAC55659.1"
/db_xref="GI:1388129"
/translation="SSFLCVWJWRSQYCFISNKYFKSVGDSVYKAVPTTINYKRRP
TLDSEFVHNHREVQRKNFLIEQGLGFINEGSGSASVEEDNFEAEQMDQ
RESEGTSTPANFVNDVAGLKIERAORSCHVSLKADRLCKYGYDKFQFSRKK
TLVSKFLNLTENAPFPMIIBADLANSVREFYLGELILFELFKLNRYSELE
VVEFRKICRAGDARLLVYRSDSLYKGRRLPKDMEYEFVFTGGLLPNDKEF
SGNGLPHANTFTAAFSIRNSYRNFIENDSCR:RLYEAPPGGKTHLIAFSVK
MHKKNRILVTANKSSQVEILKINDSKREHETKLLKFKASKAERNYPSADSIVY
TIDVYLMHLGTCVDLFDCEPMHAGAVTAVNTLVPVSFFVFGDSQ:HHIERNE
YDVSFSDLDRLVAADRLYGVGSYPCPDVCGMLSKHPNTVATTNVESEKSLTI
TEINSVDDVVASKRNTYLFLOSEKKELEKHLAKKGVKATVTVHEAQDTRDVLY
RTFQDAPPSFNHINVAITRHTESLTVAVLARRENIAAACEANSELVCFRLTP
HFGGSLNIDVEPVTYDMSRCKASSAPINSDFLYEDVVPGSTSLNFQ:SAEMESQ
PFESGANNVTVDSPKSGSGTDHDEQRV"
<2132..3514
/function="replication"
/notes="ORF1b; similar to RdRp's of other closteroviruses;
the BYSV RdRp is presumably expressed via a +1 ribosomal
frameshift"
/codon_start=1
/product="RNA replicase"
/protein_id="AAC55659.2"
/db_xref="GI:4559421"
/translation="SVVRQAIIPKRRPSQEN:LSYEARNYNFITERSGPHEYGRA
MAANVLKRCPLDKVAARISDIIITAIRGRFRELWLSKRSFSQKSLASDLETPLNLEE

```

```

IVRFKLMVKRDAXVKLDASCLVKHPPAOTIMFHRKAINALYSCQDFEFKKNRVLSCSP
NIIFTEMTNQKENDIASNLLGPNVFNIGEDISFYDKSDQDAFIKSFDCALYEERGF
GFDPFELLVWMEGETYRSERNTMDQVGTIVCORKSGASNTWIGNSVTVIGILAMYID
VSKFQALFVSDSLIPSADEIAELGFTFKFLTPSPVYFCSKFLVFTGDK
CVFVDPYKLLVKLGASGRRLSDSELFVFEVSRDLTKFEGDERVNLTLSELVHLKYE
FEMGNTKALSTIHLRSNLSFKLTVKRTGKVVYGRKAYILKKFLGYNIEPITTT
FGDAWVFYKE"
3522..4343
/function="unknown"
/notes="ORF2"
/codon_start=1
/product="p30"
/protein_id="AAC55660.1"
/db_xref="GI:1388131"
/translation="MALYKYSALALIVDPDRYTSFENLRAIFSACDARVFGVSVVC
GSVJASTJPTLTSSVFEVQNGEFLSKRYRMDERRRITITGAEMT:IPHSDAKILFHVQ
VKKO:DGNTGULLVYNGPHLIGGTIVRQFDILQGCENAYADLACPSRFSEFVSJVL
CDPMFSEIINTSDNNRSACTDAFYRRVVFESDGLLTDVDSNHRHVGLDRASE
SSRGDPRDVLSSGEPEKXVLNRN:MCLFLVLLLAGATIVVYVNPNSRFGVS"
4401..4556
/function="unknown"
/notes="ORF3; similar to small hydrophobic proteins of
other closteroviruses"
/codon_start=1
/product="p6"
/protein_id="AAC55661.1"
/db_xref="GI:1388132"
/translation="MDCILRAFUPFGPAJ:VICFFIAVAAYFFAFVKNTHSQDQTDVDI
RQEDLAG"
4574..6409
/function="unknown"
/notes="ORF4; similar to HSP70 homologs of other
closteroviruses"
/codon_start=1
/product="p66"
/protein_id="AAC55662.1"
/db_xref="GI:1388133"
/translation="MVVGVGVDFGTFSSVCVFNSGRLHVFQKQNSAYIPTCLFLYSDT
MAMSGFYDAETASLDPNVKGFGFRDLKRWVGCDE:NIIEYKSLKLPHYSVTLNFGKG
SRK:PTLGSYSGSVOMSGSLGLIALP:QALVKSAATFEKCECTELIVSVPANCDMQ
RLFTENCUNLSGFTCVHMNEPDSAAA:YQMANLPFDEEAD:SSLKESJ:K:DPITY
NQFTVRSAGDMNLGRRDVRAPAK:YQMANLPFDEEAD:SSLKESJ:K:DPITY
TVKTKDGEKTVVYVSRGL:JAEVIVFVDRTIKVMKRVFELYVKMNLKADQAKASLV
VGSYSLPGLKSLQVDFVSECIDLPDRAAAGCALYSSLSSESPPKLLVDCASH
NLS:PNVYGESIVLPAGAPVPFVGTDRDINLASCVSGSVPV:FEGDRTKCFYDKV
FSGTVPLKDLGVTGDC:PTRTIVTLATEVSVGVTKFTITG:SAKKVFVGVPAYDFSK
EVSIRSITELHTDNQNRVLLALTLTKTREARQKESYSEKHLDLSLGNLDTESKXK
FNGYNEQTADVCRILLGKSVQKTLRGARVEELSYRNI:EVQTLK1"
6285..7889
/function="unknown"
/notes="ORF5; similar to analogously positioned proteins of
other closteroviruses"
/codon_start=1
/product="p61"
/protein_id="AAC55663.1"
/db_xref="GI:1388134"
/translation="WSRRPTFAGYSWGLFKRHYGEPEKMSYLTETSMKYKPLKSESI
TFYDGSILTSAELPARSGTAIEYEAIALIFSDSITTKSEKLSIYRGLNQ:INHSIY
AETELVDYDKTICKET:ISAVESFMGRASAAQVHCMSLSNSCGELINPDTARFI
OLVPRKDAEQAOQNTSGSVSDVLYVLCQLYDMSKKSNAAGRTQLMESVSFIRDFP
QHSLEYXVSPDNQVLTGVDLYDLCIEHNRLSYLKNLDNFRJFKQTYLPMIDIFDY
SWEIYARDELLFPIDPEY:IKVPTMTVIDANVLYKNLVDLYDLY:ENKSI:LALEKK
I:ISILCRNDEIGECALWAAFPCVGYTARQVRVVRPDTYELDGI:FSKPI:VMSGVE
LFPPELOKRPVDSVLSRRFNGAKAGEAITVFKIGISFPPT:TRLNAPSKYSYLNDVF
QVANSLGITEPEKILCNIAKDUDVMCAQRISSWKAP:PIAORNGEAINSAKIRTLPTN
TLNALSBCULUNQAPSWNNTLTINLR"
7832..8515
/function="unknown"
/notes="ORF6; similar to analogous BYSV coat proteins of
other closteroviruses"
/codon_start=1
/product="p25"
/protein_id="AAC55664.1"

```

```

/db_xref="GI:1398135"
/translation="MPDPQGAELVEHNANKSSLEVFSETSEKVKCKFENNTHKTFKO
VNPILLNDELREVLCKKTELKTLKALDELIYHVAFFLLRAVSFVTSKYBYKGS
YYSIDQRYITVNDRAWFPQVILASKNKPNGLRAFCAJEGMTLSVARGPDPFGT
RVSGKRGAPSGSEYLGADFLTSTCPMLSDHRAVALSASRNALDRSAASQ-DKKXVSL
YDFGKVVT"
misc_feature
8551 ..
/notes="transcription start site for coat protein
subgenomic RNA"
CDS
8617..9264
/notes="CP; CRF7; similar to coat proteins of other
cistoviruses"
/codon_start=1
/evidence=experimental
/product="coat protein"
/protein_id="AAC55665.1"
/db_xref="GI:1398136"
/translation="VAGNDESDSSASQVTAKDMIFAPENFARASATCJNGENX
KXLFEEFSVRKTDQVTESGIPTILMTLYALATSTSKIDIEDKTLPLVSAKIDAVN
VTITVEDIKFNVSUTLLKXNKLRLVFARTEEYLRFCYKHILPNIARANKHG-
PADYSYLAADVFQTSNLLKHEQAVLLEGNAATASSGTTRESAVNLVYLGSSK"
9261..9743
/function="unknown"
/notes="ORF8"
/codon_start=1
/product="p18"
/protein_id="AAC55666.1"
/db_xref="GI:1398137"
/translation="KXVIESLQFVTPVHEKYALVLTDDAFIFELHGVQSSYPVL
WHGEAGADGSGVISTFYKELQVNLNSLQYRMLRVKRTTSSLEJINVRANAV
SIFRCNAPLIFVTSKGPESDAVLNYGIIGTPSVCPTHLSS-VESYGNITHFAE
"
9746..10306
/function="unknown"
/notes="ORF9; similar to the 20-kDa protein of CTV and to
the 21-kDa protein of BYV, two other closteroviruses"
/codon_start=1
/product="p22"
/protein_id="AAC55667.1"
/db_xref="GI:1398138"
/translation="WKLLSDSYVYDSINLANKNELNKNKSVPLESIASYNELLK
VNAQSGDAIRKNEKNENLFPDLSVNTDLEMLVRDILRAEINKFNCSSQDQV
KMIIRINTHFSNRSRERPEVTEQSFMSRQVINTSLQNSYKYFEVTINYPKCA
LLEFVXNGSVKTFVENVTESEF"
10304..10545
/notes="similar to 3' UTR's of CTV and BYV"
3'UTR
2989 a 2196 c 2497 g 2863 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 6,95e-18 Length: 10545
Score: 266.50 Matches: 66
Percent Similarity: 48.09% Conservative: 22
Best Local Similarity: 36.07% Mismatches: 94
Query Match: 26.89% Indels: 1
DB: 14 Gaps: 1
US-09-613-486-15 (1-198) x BYU51931 (1-10545)
QY 11 AsnLeuValIleThrAspAlaSerSerLeuAsnGluValAspLysLysLeuSerAla 30
Db 8704 ANTTTCGTCGCGAGCGGACCTTTGTAATGGTGAAGAAACAAAGAGTGTTCGAG 8763
QY 31 GluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGluValValPhe 50
Db 8764 GAATTTTCTGTGAGAGTGAAGACTCAGAGTGTCTAGTGAATCCGCATCCCACTTTG 8823
QY 51 GlyLeuLeuLeuTyralaLeuAlaAlaArgThrThrSerProLysValGlnArgAlaAsp 70
Db 8824 GGAATGACGCTGTACCCCTCTAGTACCTTAAGCACTTCTCTAAATTCGATATCGAGAT 8883
QY 71 SerAspValIlePheSerAsnSerPheGlyGluArgAsnValValThrGluGlyAsp 90
Db 71 SerAspValIlePheSerAsnSerPheGlyGluArgAsnValValThrGluGlyAsp 90

```

```

Db 8884 AAGACACCGCTAGTAGCGGCTAAGATAGATGCGGTCAACGTCAGCATAAATTACGAGAC 8943
QY 91 LeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuAlaGThr 110
Db 8944 ATTAAAGAACTTCGTAAACACTCCITGACGTTCGTTAAGAACTATAAGAAATAAACTCCAGTC 9003
QY 111 PheGlyArgThrPheThrGluAlaTyValAspPheCysIleAlaTyLysHisLysLeu 130
Db 9004 TTGCTCGGACTTTTGAGGAGGAGTACCTACGCTTGTGAGGCAATACAAACACATCCCTC 9063
QY 131 ProGlnLeuAsnAlaAlaGluLeuGlyLeProAlaGluAspSerTyLysLeuAlaAla 150
Db 9064 CCGAATATCGAAGAGCGAACAACACGCGTATCCCGCTGATCTACTCGATTATGACGCT 9123
QY 151 AspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMetPheAla 170
Db 9124 GACTTCGTCCAACTAGCAACTTATTGAAAGAGCAGCAAGGCA---GTACTTCTGGAA 9180
QY 171 SerMetTyralaLeuLysThrGluGlyGlyValValAlaThrProValSerAsnLeuAla 190
Db 9181 GGTAGAAACGCTCTACGATCTTCGCGCACCACTAGGAGTCTCGGTCACACTTAAAA 9240
QY 191 GlnLeuGly 193
Db 9241 TATCTTGGG 9249
RESULT 15
AF501867
LOCUS
DEFINITION Citrus tristeza virus coat protein gene, complete cds.
ACCESSION AF501867
VERSION AF501867.1 GI:20454248
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 672)
Virus; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
AUTHORS
Roy, A., Ramachandran, P. and Brilansky, R.H.
TITLE
Grouping and comparison of Indian citrus tristeza virus isolates
based on coat protein gene sequences and restriction analysis
patterns
JOURNAL Arch. Virol. 148 (4), 707-722 (2003)
MEDLINE 22548851
PubMed 12664295
REFERENCE
2 (bases 1 to 672)
Roy, A., Ramachandran, P. and Brilansky, R.H.
Direct Submission
Submitted (12-APR-2002) Plant Pathology, University of Florida, 700
Experiment Station Road, CREC, Lake Alfred, FL 33850, USA
FEATURES
source
1..672
/organism="Citrus tristeza virus"
/mol_type="genomic RNA"
/db_xref="taxon:12162"
/country="India: Bangalore"
1..672
/codon_start=1
/product="coat protein"
/protein_id="AA022216.1"
/db_xref="GI:20454249"
/translation="MDDETLLKNNKNNKPKBGGDVVAESSFGSNLHIDFLIAMND
VRQGTQQAALNRDLTKGKYNLPDEKDPH:AMMLYRLAKVSSLSQSDDTTG
ITTRREGVEVDSKLTWTFNSKIGTNRNALRVGRTNDALVAFRCRNRLSYG
GRPLDAGIPAGYHYLCADFLTGAGLTGECVVIQAKEQLLKRGADEVVTVNRQLG
KFNR"
BASE COUNT 191 a 124 c 177 g 180 t
ORIGIN
Alignment Scores:
Pred. No.: 7.71e-05 Length: 672
Score: 130.50 Matches: 61
Percent Similarity: 40.00% Conservative: 27

```

Best Local Similarity: 27.73% Mismatches: 75  
Query Match: 13.17% Indels: 57  
DB: 14 Gaps: 11

JS-09-613-486-15 (1-198) x AF501867 (1-672)

QY 1 MetGluLeuMetSerAspSerAsnLeuValIleThrAspAlaSerSerLeu 20  
DB 91 ATGAACCTTACATCGATCGGACTCTG-----ATAGCGATGAACGATGTGGTCAGTTG 144  
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40  
DB 145 GGAACCCCAACAGACCCGCTTTGAACAGAGATTG-----TTTCTTACTCTGAAGGCG 198  
QY 41 -----AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAla 56  
DB 199 AAGTATCCTTAACCTTGCCTCAGCAGGATAAGGACTTCCACATAGCTATGATGTATATCTCT 258  
QY 57 LeuAlaAlaArgThrThrSer----- 63  
DB 259 TTAGCGGTTAAGAGTTTCATCGTTGCAAGTGATGACACACCGGSCATAACATAACT 318  
QY 64 -----ProLysValGlnArgAlaAsp-----SerAspValIlePhe---Ser 76  
DB 319 CGGGAGGGTGTGAAAGTGGATTGTCTGACAAAGCTTTGGACTGACGTCGTGTTAACTCC 373  
QY 77 AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLysValLeuAsp 96  
DB 379 AAGGTATTGGTAACCGT----- 396  
QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116  
DB 397 -----ACTAATGCCCTTCGAGTCTGGGCTAGGACTAACGAT 432  
QY 117 GluAlaTyrValAspPheCysLeuAlaTyrLysHisLysLeuProGlnLeuAsnAlaAla 136  
DB 433 GCCCTTTATTAGCTTCTGT---AGACAGAAATCGCAATTGAGTTATGGTGGACGTCGC 489  
QY 137 AlaGluLeuGlyIleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCys 156  
DB 490 CTAGATCGCAGGGGATTCGGCTGGATATCATATTACCTATGTCCAGATTTCCTG---ACCGA 546  
QY 157 ProLysLeuSerGluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLys 176  
DB 547 GCTGGCTTGACTGATTAGNAATCGCTGTGTACATACAAGCTAAGGAACAATTTGTGAG 606  
QY 177 ThrGluGlyGly-----ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArg 194  
DB 607 AAGCGAGGGGCTGATGAAGTCGTA-----GTTACTAATGTCAGGAGCTTGGGAAA 657

Search completed: November 8, 2003, 00:59:54  
Job time : 3533 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OX protein - nucleic search, using frame\_plus\_p2n model

Run on: November 7, 2003, 23:40:55 ; Search time 267 Seconds  
(without alignments)  
2001.830 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMSDLSNLVITDASSL.....GVVNTPVSNLRQLGRREV 198

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5135512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DBV=xih  
-Q=/cgn2\_1/USPTO.spool/US09613486/runat\_07112003\_120411\_27082/app\_query.fasta\_1.391  
-DB=N Geneseg 19Jun03 -QFMT=fastap -SUFFIX=ring -M.NMA=CH=0.1 -DOFCS=0  
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=bicsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09613486.acgn\_1\_1\_312.grunat\_07112003\_120411\_27082 -NCPU=6 -ICPU3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=2.0 -XGAPEXT=0.5 -FGAPEOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1930.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	597	20	Grapevine leafroll
2	991	100.0	15500	20	Grapevine leafroll
3	286	29.9	688	16	Sugar beet yellows
4	285	28.8	783	16	BYV coat protein c
5	89.5	9.0	1500	24	C. albicans BAX-as
6	86	8.7	1197	21	Arabidopsis thaliana
7	86	8.7	1199	21	Arabidopsis thaliana
8	86	8.7	349980	22	Pyrococcus abyssi
9	85	8.6	4467	18	PSR0800 fragment 1
10	84.5	8.5	4766	22	GLRAV-5 nucleotide
11	83.5	8.4	640681	24	Buchnera sp. genom
12	81.5	8.2	1209	21	Erysipelothrix rhu
13	81.5	8.2	1881	21	Erysipelothrix rhu
14	81.5	8.2	204C	21	Erysipelothrix rhu
15	81.5	8.2	2814	21	Erysipelothrix rhu
16	81.5	8.2	910715	20	Borrelia burgdorferi
17	81	8.2	349980	24	Bifidobacterium lo
18	80.5	8.1	10709	22	Genomic fragment #
19	80	8.1	2133	24	Streptococcus poly
20	80	8.1	2187	24	Streptococcus poly
21	80	8.1	2365369	24	Genomic sequence o
22	79	8.0	1356	23	Enterococcus faeca
23	79	8.0	1434	18	Grapevine leafroll
24	79	8.0	2572	23	Drosophila melanog
25	79	8.0	4827	23	Drosophila melanog
26	79	8.0	11597	20	Enterococcus faeca
27	79	8.0	11597	21	Enterococcus faeca
28	79	8.0	17915	24	Grapevine leafroll
29	79	8.0	6313	23	Propionibacterium
30	79	8.0	534720	19	Rhizobium species
31	79	8.0	536165	19	Rhizobium species
32	78.5	7.9	2282	22	Nucleotide sequenc
33	77.5	7.8	665	24	Bacillus clausii 9
34	77.5	7.8	4494	23	Drosophila melanog
35	77	7.8	912	22	Aspergillus nidula
36	77	7.8	929	22	Aspergillus nidula
37	77	7.8	2091	24	Candida albicans e
38	76.5	7.7	7676	19	Human galactokinase
39	76.5	7.7	8095	24	Gene #228 used to
40	76	7.7	600	22	Human foetal liver
41	76	7.7	600	22	Probe #8204 for ge
42	76	7.7	600	22	Human brain expres
43	76	7.7	600	22	Human bone marrow
44	76	7.7	600	22	Probe #7406 for ge
45	76	7.7	600	22	Probe #1068 used

ALIGNMENTS

RESULT 1  
AAV08870  
ID AAV08870 standard; cDNA; 597 BP.  
XX AAV08870;  
AC AAV08870;  
XX  
XX  
XX 29-MAR-1999 (first entry)  
XX  
XX Grapevine leafroll virus type 2 coat protein ORF6 product.  
DE  
XX GRVAV-2; closterovirus; grape; tobacco; transgenic plant;  
KW disease resistance; virus resistance; beet yellows virus;  
KW tristezza virus; coat protein; ss.  
XX  
XX Grapevine leafroll virus type 2.  
OS  
XX WO9853055-A1.  
PN

26-NOV-1998.  
 20-MAY-1998; 98WC-US0313.  
 20-MAY-1997; 9TUS-CC47194.  
 (CORR ) CORNELL RES FOUND INC.  
 Gonsalves D, Ling K, Zhu H;  
 WPI; 1999-0451C7/04.  
 P-9SDB; AAW73482.  
 Grapevine leafroll virus (type 2) proteins and polypeptides - and  
 encoding DNA, useful e.g. to impart grapevine leafroll resistance to  
 grape and tobacco plants and detect grapevine leafroll virus  
 Claim 38; Page 44; 151pp; English.  
 This is the nucleotide sequence of open reading frame ORF6 of  
 grapevine leafroll virus type 2 (GLRAV-2) RNA (see AAV08874). It  
 codes for a 22 kDa coat protein (see AAW73492). The GLRAV-2  
 genome includes 9 open reading frames (see AAV08954-72) for a  
 polyprotein, an RNA-dependent RNA polymerase, heat shock proteins,  
 coat proteins and proteins of unknown function (see AAW73476-84).  
 These can be used to produce antibodies useful for detecting  
 GLRAV in samples e.g. by ELISA (claimed). The nucleic acid  
 molecules can be used to produce probes and primers for such  
 detection, and to transform host cells (especially Agrobacterium  
 cells), Agrobacterium tumefaciens, grape, citrus, beet or tobacco  
 cells) and produce transgenic plants (claimed). They can be used  
 to impart GLRAV-2 resistance to Vitis scion or rootstock cultivars  
 or Nicotiana cultivars (claimed). Because extensive similarity  
 exists between hsp70-related sequence regions of GLRAV-2 and other  
 closteroviruses, the DNA may also be used to impart beet yellows  
 virus resistance to beet cultivars or tristeza virus resistance to  
 citrus scion cultivar/rootstock cultivars (claimed).  
 Sequence 597 BP; 165 A; 121 C; 157 G; 154 T; 3 other;

301 CTCACGAGGTTCACTAATAAACTTACACGTTGGTGGTACTTCTCAGCTAGGCTTACGTT 360  
 121 AspPheCysIleAlaTyrIlyshisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
 361 GACTTTTGATCGCGTATAAGCACAAATTACCCCAACTCAACGCGCGCGGAATTGGGG 420  
 141 IlePrcAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 421 ATTCAGCTGAAGATTCGTTACTAGCTGCAGATTTCCTGGTACTTGCCTGAAGCTCTCT 480  
 161 GluLeuGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180  
 481 GAATACACCAAGTAGGAAGATGTTCCGAGTAGTACGCTCTAAAACTCAAGGTGGA 540  
 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgGluValMet 198  
 541 GTGGTAATACACCAAGTAGGAAGATGTTCCGAGTAGTACGCTCTAAAACTCAAGGTGGA 594  
 RESULT 2  
 AAV08874  
 ID AAV08874 standard; cDNA; 15500 BP.  
 AC AAV08874;  
 XX 29-MAR-1999 (first entry)  
 DE Grapevine leafroll virus type 2 (GLRAV-2) genome.  
 XX  
 KW GLRAV-2; closterovirus; grape; tobacco; transgenic plant;  
 KW disease resistance; virus resistance; beet yellows virus;  
 KW tristeza virus; protease; methyltransferase; helicase;  
 KW heat shock protein; coat protein; RNA polymerase; ss.  
 XX Grapevine leafroll virus type 2.  
 OS  
 XX  
 PH Key  
 FT CDS Location/Qualifiers  
 FT 4..7923  
 FT /tag= a  
 FT /product= polyprotein (protease, methyltransferase,  
 FT /note= "Claim 26"  
 FT 7922..9301  
 FT /tag= b  
 FT /product= RNA-dependent RNA polymerase  
 FT /note= "Claim 29"  
 FT 9365..9535  
 FT /tag= c  
 FT /product= hydrophobic protein  
 FT /note= "Claim 43"  
 FT 9551..11350  
 FT /tag= d  
 FT /product= heat shock 70 protein  
 FT /note= "Claim 32"  
 FT 11277..12932  
 FT /tag= e  
 FT /product= heat shock 90 protein  
 FT /note= "Claim 35"  
 FT 12844..13515  
 FT /tag= f  
 FT /product= diverged coat protein  
 FT /note= "Claim 41"  
 FT 13584..14180  
 FT /tag= g  
 FT /product= coat protein  
 FT /note= "Claim 38"  
 FT 14180..14665  
 FT /tag= h  
 FT /product= undefined  
 FT /note= "Claim 45"  
 FT 14667..15284  
 FT /tag= i  
 FT /product= undefined  
 FT /note= "Claim 47"  
 FT

3'UTR 5285...15500  
 /tag= 3  
 /notes= "Claim 48"  
 WO9853055-A1.  
 26-NOV-1998.  
 20-MAY-1998; 98WO-US10313.  
 20-MAY-1997; 97US-0047194.  
 (CORR ) CORNELL RES FOUND INC.  
 Gonsalves D, Ling K, Zhu H;  
 WPI; 1999-045307/04.  
 P-PSDB; AAW73476, AAW73477, AAW73478, AAW73479, AAW73480, AAW73481,  
 AAW73482, AAW73483, AAW73484.  
 Grapevine leafroll virus (type 2) proteins and polypeptides - and  
 encoding DNA, useful, e.g. to impart grapevine leafroll resistance to  
 grape and tobacco plants and detect grapevine leafroll virus  
 Disclosure; Page 12-20; 151pp; English.  
 A total of 15,500 bp of the RNA genome of grapevine leafroll virus  
 type 2 (GLRAV-2) has been sequenced and cloned from GLRAV-2  
 isolated from infected Vitis vinifera cv. Pinot Noir. About 85%  
 of the total RNA genome was revealed from 2 different clones. The  
 sequence in the coat protein gene region was determined and  
 confirmed from several overlapping clones. The genomic  
 organisation of GLRAV-2 includes 9 open reading frames (see also  
 AAV08864-72) encoding a polyprotein (see AAW73476) having papain-like  
 protease, methyltransferase and helicase motifs, an RNA-dependent  
 RNA polymerase, heat shock proteins, coat proteins and other  
 proteins of unknown function (see AAW73476-84). These proteins can  
 be used to produce antibodies, useful to detect GLRAV in samples  
 e.g. by ELISA (claimed). The nucleic acid molecules can be used to  
 produce probes and primers for such detection, and to transform  
 host cells (especially Agrobacterium vitis, Agrobacterium  
 tumefaciens, grape, citrus, beet or tobacco cells; and produce  
 transgenic plants (claimed). In particular, they can be used to  
 impart GLRAV-2 resistance to Vitis scion or rootstock cultivars or  
 Nicotiana (claimed). Because extensive similarity exists between  
 hsp70-related sequence regions of GLRAV-2 and other closteroviruses,  
 the DNA may also be used to impart beet yellows virus resistance to  
 beet cultivars or tristezza virus resistance to citrus scion  
 cultivar/rootstock cultivars (claimed).  
 SQ Sequence 15500 BP; 3984 A; 3178 C; 3970 G; 4368 T; 0 other;

Alignment Scores:  
 Pred. No.: 8,13e-115 Length: 15500  
 Score: 991.00 Matches: 198  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-613-486-15 (1-196) x AAV08874 (1-15500)

Qy 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerLeu 20  
 DR |  
 DR |  
 Db 13584 ATGAGTTGATG:CCGACACACCTTAGCACTGGTGATACCGACGCTTAGCTTA 13643  
 Qy 21 AsnGlyValAspIlyLysLeuLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40  
 Db 13644 AATGGTGTGCACAGAGCTTTTATCTGCTGAGTTGAAAATTTGGTGCAAGAGG 13703  
 Qy 41 AlaProhncGluGlyIleGluValValPheGlyLeuLeuLeuTyraLeuAlaAaAG 60  
 Db 13704 GTCCTAACGAGGGTATAGAAGTGGTGTTCGGTCTACTCCTTTACGCACTCGCGGAAGA 13763

Qy 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80  
 Db 13764 ACCACGTCTCTAAGGTTTCAGCGCGCAGATTCAGCGTTATATTTTCAATAGTTTCGGA 13823  
 Qy 81 GluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100  
 Db 13824 GAGAGGAATGTGGTAGTAACAGAGGGTGACCTTTAAGAAGG:ACTCGACGGGTGTGGCCT 13883  
 Qy 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyVal 120  
 Db 13884 CTCAC:TAGGTTCACTAATAAATAGAACGTTCCGTCGTACTTTCAC:TAGGCTTACGTT 13943  
 Qy 121 AspPheCysIleAlaTyLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
 Db 13944 GAC:TTTGTATCGGTATAAGCACAAATTACCCCACTCAACCGCGCGGAATTGGGG 14003  
 Qy 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 Db 14004 ATTCAGCTGAGAGTTCTACTTAGCTGCAGATTTTCTGGGTACTTGC:CGAAGCTCTCT 14063  
 Qy 161 GluLeuGlnInSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGly 180  
 Db 14064 GAATTACAGCAAGTAGCAAGATGTTCCGAGTATGTACCTCTATAAAACTGAAGGTGGA 14123  
 Qy 181 ValValAsnThrProValSerAsnLeuArgGluLeuGlyArgArgGluValMet 198  
 Db 14124 GTGGTAATATACACCGATGAGCAATCTCGCTCAGCTAGTAGAGGGAAGTTATG 14177  
 RESULT 3  
 AAG87853  
 ID AAG87853 standard; cDNA; 688 BP.  
 AC AAG87853;  
 XX  
 D: 25-MAR-2003 (updated)  
 D: 02-NOV-1995 (first entry)  
 XX  
 DE Sugar beet yellows virus capsid protein coding gene.  
 XX  
 XX Sugar beet yellows virus capsid protein; transgenic plant; ss.  
 XX  
 XX Sugar beet yellows virus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 74..688  
 FT /\*tag= a  
 FT /product= sugar beet yellows virus capsid protein.  
 XX  
 PN RJ2217820-C1.  
 XX  
 XX 15-AUG-1994.  
 XX  
 XX 27-JUN-1991; 91SU-4950054.  
 XX  
 XX 27-JUN-1991; 91SU-4950054.  
 XX  
 XX (BIOT=) BICVRCHN INST CO LTD.  
 XX (IMMU=) IMMUNOB:OTECHN INST.  
 XX  
 PI Arganovskii AA, Boiko VP, Karasev AV;  
 XX  
 DR WPI; 1995-113715/15.  
 DR P-PSDB; AAR72682.  
 XX  
 PT Sugar beet yellows virus cDNA fragment encoding capsid protein -  
 useful for production of virus-resistant transgenic plants  
 XX  
 PS Claim 1; Column 7-10; 5pp; Russian.  
 XX  
 CC The nucleotide sequence of the novel gene encoding the sugar beet yellows  
 virus (SBYV) capsid protein. The gene encodes a protein of 204 amino  
 acids with mol. wt. 22.2 kD. The gene was obtained from reverse



CC transcribed RNA isolated from purified SBV. The fragment is useful for  
 CC the production of virus resistant transgenic plants by genetic  
 CC engineering methods.  
 CC (Updated on 23-MAR-2003 to correct PN field.)  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 688 BP; 190 A; 155 C; 157 G; 186 T; 0 other;

Alignment Scores:  
 Pred. No.: 9.9e-27 Length: 688  
 Score: 286.00 Matches: 71  
 Percent Similarity: 52.91% Conservative: 29  
 Best Local Similarity: 37.57% Mismatches: 87  
 Query Match: 28.86% Indels: 2  
 DB: 16 Gaps: 2

US-09-613-486-15 (1-198) x AAQ87853 (1-688)

QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26  
 DB 104 GCGACTTTTGAAACGTAAGTCTCGCAGACCAACACTGTTTCACGGAGACACTGGCAT 163  
 QY 27 LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46  
 DB 164 AAACCTTAGGAAGAACTTCGAAGAGTGTTCGAATTAATAAGGGGTTCCGGAAGATAACCTC 223  
 QY 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrSerProLysVal 66  
 DB 224 GGAATCGCGTTAGGACTTGTGTGATTCTCTGCTACGATAGGCACCTCCAAACAAGTT 283  
 QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 95  
 DB 284 AAGCTCAACCGACGCTACCTTCATCAAGCTTCGTTTGGTGGTGGGAAGCACTGTAC 343  
 QY 86 ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105  
 DB 344 CTCACCTCAGCGTGAATTAATTCCTTTCTGGGGTCTCCAAAACCTTTTGGAGGGAACCT 403  
 QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAla 125  
 DB 404 AACAAATTCGGTGTTCCTCGGTACTTTTCAGAGGACTATATCTCTTGGCGCAGGAA 463  
 QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145  
 DB 464 TACCGAGGGAATTAACCTCCGA\*TGCCAGAGCTAACCGTCAGCGCTACCCGCTCAAGAT 523  
 QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSer 165  
 DB 524 CACTACTTAGCCGCTGACTTCATATCGAGCTGACGGAACTCACGACCTCAACAAGT 583  
 QY 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrPro 185  
 DB 584 CGTCTCTGTATAGCGCGCGGAAACGCCACTCACAGGAATTCCTCG---TCTGAATCACCG 640  
 QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194  
 DB 641 GAACCAAGTTGAACAACACTAGTCTGT 667

# RESULT 4

ID AA42975 standard; cDNA; 783 BP.

XX

AC

XX

DT

DE

XX

KW

KW

KW

KW

KW

KW

KW

XX Beet yellow virus.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 Ff sig\_peptide 13.78  
 Ff /tag= a  
 Ff /note= "Omega sequence from TMV"  
 Ff 89..703  
 Ff /tag= b  
 Ff /product= BYV\_coat\_protein  
 FT  
 XX W09428147-A1.  
 XX 08-DEC-1994.  
 XX 01-JUN-1994; 94MO-EP02786.  
 XX 02-JUN-1993; 93GB-0311332.  
 XX (SANO ) SANDOZ LTD.  
 XX (SANO ) SANDOZ PATENT GMBH.  
 XX (SANO ) SANDOZ-ERFINDUNGEN VERW GES MBH.  
 XX Boisen K, Brunstedt J;  
 XX WP; 1995-022818/03.  
 XX Recombinant DNA sequence comprising at least two coat protein genes  
 PT - used to combat viruses in e.g. sugar beet  
 XX  
 PS Disclosure; Page 11-12; 23pp; English.  
 CC The sequences given in AA42975-77 represent cDNA sequences encoding  
 CC coat proteins from beet yellow virus (BYV), beet western yellow  
 CC virus, (BWV) and beet necrotic yellow vein virus (BNYVV)  
 CC respectively. These sequences are used in the recombinant DNA  
 CC sequence of the invention. This comprises nucleotide sequences  
 CC encoding the coat proteins of at least one luteovirus and at least  
 CC one closterovirus. This recombinant DNA may be used to transform  
 CC sugar beet for combating viral infection. Other plants may also  
 CC be transformed e.g. fruit such as mangoes, apples, pears, bananas,  
 CC and field crops such as sunflower, wheat, barley, maize, and  
 CC vegetables such as potatoes, carrots, cabbage and onion.  
 XX  
 SQ Sequence 783 BP; 233 A; 178 C; 171 G; 211 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.6e-26 Length: 783  
 Score: 285.00 Matches: 70  
 Percent Similarity: 53.44% Conservative: 31  
 Best Local Similarity: 37.04% Mismatches: 86  
 Query Match: 28.76% Indels: 2  
 DB: 16 Gaps: 2

US-09-613-486-15 (1-198) x AA42975 (1-783)

QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26  
 DB 104 GCGACTTTTGAAACGTAAGTCTCGCAGACCAACACTGTTTCACGGTGAAGACTCGAC 178  
 QY 27 LeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIle 46  
 DB 179 AAACCTAGGAAGAAATTCGAAGAGTGTTCGAATTAATAAGGGGTTCCGGAAGACAACATC 238  
 QY 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrSerProLysVal 66  
 DB 239 GGTCTCGCGTTAGAACCTTGTGTGATTCTCTGCGACGATAGTACTTCTATAAAGTT 298  
 QY 67 GlnArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85  
 DB 299 AGTCTCCAAACCGACGCTACTTTCATCAACAGCTTCGTTTCGTCGTGGGAAGAAATTGTC 358  
 QY 86 ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105

```

Db 359 CTCACTCAGGTGAAGTGAAGTCTTTCTGGACTCTCAGAACTTTTGGGAAGAAAGCTT 418
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAa 125
Db 419 AACAAAGTTGGTTGTTCTCCGACACTTTCAGAGGACTACATATCCTTCGCGAAGGAA 478
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyLeuProAlaGluAsp 145
Db 479 TACCGAGGAAGACTGCTCCGATGTGTAGAGCAACCGTCACGGTCTACTGCTGAAGAT 538
QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 539 CACTACTTACTGCTGATTTTCATATCAGATCAACAGAACTTACTGACCTACACAAAGT 598
QY 166 ArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGlyValValAsnThrPro 185
Db 599 CGTCTGCTGTGTGGCGGCGGAAACGGCCACTCACACAGAGTTCTCG---TCTGAATCACCA 655
QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194
Db 656 GTAACAGTTTGAACAGC---GGGTGCT 682

RESULT 5
ABQ76539
ID ABQ76539 standard; cDNA; 1500 BP.
XX
AC ABQ76539;
XX
XX 21-NOV-2002 (first entry);
XX
DE C. albicans BAX-associated cDNA fragment SEQ ID 503.
XX
XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
XX vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
XX apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
XX neurodegeneration; cell death; ss.
XX
XX Candida albicans.
XX
XX W200264766-A2.
XX
XX 22-AUG-2002.
XX
XX 21-DEC-2001; 2001WO-EPI5398.
XX
XX 22-DEC-2000; 2000EP-0870318.
XX
XX 04-JAN-2001; 2001EP-0870002.
XX
XX 09-JAN-2001; 2001EP-0870003.
XX
XX (JANC ) JANSEN PHARM NV.
XX
XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
XX WPI; 2002-667002/71.
XX
XX P-PSDB; ABG93273.
XX
XX New isolated nucleic acid representing a synthetic Bax gene, useful as
XX medicament for treating, preventing and/or alleviating yeast or fungal
XX infections or proliferative disorders, or for preventing apoptosis in
XX certain diseases.
XX
XX Claim 36; Figure 2; 344pp; English.
XX
XX This invention describes a novel nucleic acid representing a synthetic
XX Bax gene. The Bax gene of the invention is useful for identifying
XX Bax-resistant yeast or fungi, identifying, or obtaining and identifying
XX Candida spp. sequences that are differentially expressed in a pathway
XX eventually leading to programmed cell death or identifying inhibitors or
XX inhibitor sequences of Bax-induced cell death. The products of the
XX invention have cytostatic, fungicide, immunosuppressive, virucide and
XX vasotropic activity and can be used in vaccines or for gene therapy. The
XX isolated nucleic acids, polypeptides, pharmaceutical compositions,

```

```

CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenic flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polynucleotide associated with the Bax gene
CC described in the disclosure of the invention.
XX
SQ Sequence 1500 BP; 442 A; 291 C; 309 G; 458 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.387 Length: 1500
Score: 89.50 Matches: 53
Percent Similarity: 43.60% Conservative: 39
Best Local Similarity: 25.12% Wismatches: 76
Query Match: 9.03% Indels: 43
DB: 24 Gaps: 12
XX
US-09-613-486-15 (1-198) x ABQ76539 (1-1500)
QY 3 LeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSer---LeuAsn 21
Db 517 TTGATGGCTTCTGGAAATGGGACAGCTTCTGCTACTGCTTCTACCACTGTTTGAAG 576
QY 22 GlyValAspLysLysLeuLeuSerAla---GluValGluLysMetLeuValGlnLysGly 40
Db 577 ACTGCTGAATCCACCCCATATCTGCTTATCTTCCCAATTGTAGTCAAGCCGGT 636
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg 60
Db 637 ATGCCAAAGGTGTTTCAACATTTGTTTCGGTTTGGTGTCTGCTGCTGCTGCTGCTGCT 696
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 697 GCTAACATCCAAAGATTGAAAAA-----GTGCTTTTCACTGGTTCTACTGCC 744
QY 81 GluArgAsnValValVal-----ThrGlyAspLeuLysLysVal-----94
Db 745 ACCGGTAAATTTATCATGAATTCGGTGTGCTGAATCAAACTTGAAAAAAGTTACTTTGGAA 804
QY 95 LeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThr 114
Db 805 TTGGTGGTAAATCTCCAAACATTTTTCACCAT-----GCTGATTTCGACAGACT 858
QY 115 PheThrGluAlaTyrVal-----AspPheCysIleAla-----125
Db 859 ATTCAAACTTGATTGTTTCTATCTTACAAATCTGCTGAAGTCTGTTGCTGGTTCT 918
QY 126 -----TyrLysHisLysLeuProGlnLeuAsnAlaAla 137
Db 919 CGTCTTTTGAATTCGCGTGTTCAGACCAAGTGTGTGAAAAAATTCAAAGAGAGCTGCT 978
QY 138 Glu-----LeuGlyIlePro---AlaGluAspSerTyrLeuAlaAaAspPheLeu 153
Db 979 GAAAGTGTCAAGGTGTGTAACTTCAGCAAGACACTTTTATGGGTGCC-----1029
QY 154 GlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMetPheAlaSerMetTyr 173
Db 1030 -----CAAGTTTCTGAGCTCAATTTGTC-----AAAATTTTGAATACGTTGAA 1074
QY 174 AlaLeuLysThrGluGlyGlyValValAsnThr 184
Db 1075 TCTGTTAAATCTCAAGGTGCTACTGTTGTTTACC 1107
XX
XX RESULT 6
XX AAC45964
XX ID AAC45964 standard; DNA; 1197 BP.
XX
XX AAC45964;

```

XX 18-OCT-2000 (first entry;  
DI Arabidopsis thaliana DNA fragment SEQ ID NO: 48418.  
XX DE  
XX DE  
XX DE  
XX Hybridisation assay; Genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; Promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
PN SP2033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX  
PF 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134220.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145929.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148119.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.

```

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161520.
PR 28-OCT-1999; 99US-0161592.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Alignment Scores:
Pred. No.: 0.78
Score: 86.00
Percent Similarity: 38.92%
Best Local Similarity: 26.49%
Query Match: 8.68%
DB: 21
Length: 1197
Matches: 49
Conservative: 23
Mismatches: 93
Indels: 20
Gaps: 9

```

US-09-613-486-15 (1-198) x AAC45964 (1-1197)

```

Qy 7 SerAsnLeuSerLeuValIleThrAspAlaSerSerLeuAsnGlyVal-----Asp 24
Db 427 TCGGAACCAACCAACAGTTTCAAGAGTTACATTGCTGCTCAAGGTATCGTGGAGAC 486
Qy 25 LysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGlu 44
Db 487 AAGAAACGATCGAGATCATAAACATGCCTTCGTGGTTGTGAGTCAGGGCCCTACCAT 546
Qy 45 GlyTle-----GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg 60
Db 547 TTCAATTGAACTATTACGAATTCCTCAAGGCGTCTCGAGTATCCTTTCTTCTGCT 606

```

```

Qy 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 607 TACCAAGACTTTATCTTAAGAGCGCTTGAAAAATTCCTGGGAGGCTTTACAGTTTAGT 666
Qy 91 GluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100
Db 667 GTCCGGATGTTTGTT-----GGAGGTTTACCGCAATG-----GGTGTTTACCG 714
Qy 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120
Db 715 ATC---CATATGACTGCTAAATTCGCAACATTTTAGATTCTGCTTGAACACCATTAAC 771
Qy 121 AspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
Db 772 AAGACTCTGTTTATACAAATGAGAAACTCAGAACTCTTGCCCCCAATCGAAGCATCT 831
Qy 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
Db 832 CTTCACGGA---ACCAATTCCTTTACGCCGATGCTCTATAAT-----CCTATGATGGAG 882
Qy 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLys---ThrGluGly 179
Db 883 ATGATCCAAAACCTACGAAA-----TACGGGTTCAAGAGACGACGAGAGA 927
Qy 180 GlyValValAsnThr 184
Db 928 GGATGTTGTGGACA 942

```

RESULT 7  
AAC39C12  
ID AAC39C12 standard; DNA; 1199 BP.  
XX AAC39C12;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23034.  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134258.  
PR 14-MAY-1999; 99US-0134259.  
PR 14-MAY-1999; 99US-0134261.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139219.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-014287.  
PR 01-JUL-1999; 99US-0142842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142190.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0143920.  
PR 12-JUL-1999; 99US-0143977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147018.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147535.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.

XW	Hyperthermophilic archaeon.; hyperthermophilic protein.; ds.
XX	OS
XX	Pyrococcus abyssi.
XX	Location/Qualifiers
PH	Key
FT	misc_feature
FT	300001...349980
FT	/tag= "a
FT	/note= "This sequence overlaps with the 5'
FT	AAH41223"
XX	XX
XX	FR2792651-A..
XX	PN
XX	XX

XX	
XX	
PF	22-APR-1999; 99FR-0005034.
XX	
XX	
PR	22-APR-1999; 99FR-0005034.
XX	
XX	(CNRS) CNRS CENT NAT RECH SCI.
PA	(IPRE-) IPREME INST RECH EXPL MER.
PA	

Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;

Qy	7	SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyVal-----Asp	24	XX	Accession O/ NCBI/Genbank O/ Genbank W/ NCBI/Genbank X/
				XX	
				DR	
				XX	
Db	429	TCGGAAACACCAAAATGTCAAAGGTACATGTGCTCTCAAAGGATCGTGGGAGAC	488	XX	WPI; 2001-26236/14.

QY  
25 LysLysLeuLeuSerAlaGluValGlnLysMetLeuValGlnLysGlyAlaProAsnGlu 44  
|\_|\_|\_| :|\_|\_| :|\_|\_| :|\_|\_| :|\_|\_| :|\_|\_| :|\_|\_| :|\_|\_| :  
PT New nucleotide sequences isolated from Pyrococcus abyssi encode  
P: proteins useful in industry -  
XX

DB	489	AGAAAGCGATCGAGATCATAAACATGCTTCGTGGTTGTGAGCGAGGCGCTAACCAT	54#	PS	Claim 1; Page 183-279; 1657pp; French.
QY	45	Glylle-----GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg	6C	XX	The present invention relates to the genomic sequence of <i>Pyrococcus</i>

[illegible][illegible][illegible]

Qy  
101 LeuThrArgPheThrAsnLysIleuA<sup>g</sup>gThrPheGlyA<sup>g</sup>gThrPheT<sup>r</sup>GluAlaTyrVal 129  
:::  
:::  
:::  
Sg Sequence 349980 BP; 95167 A; 78444 C; 78870 G; 97499 T; C other;  
XX  
XX

	Alignment Scores:	2,99e+03	34996C
	Pred. No.:	86.00	Matches: 40
	Score:		
717 ATC---CATATGACTGCTAAATTCGCAACATTTTATAGATTCTGCTGGACACACCATTAAC	773		
Db			
121 AspPheCysIleAlaTyrIleLysHisIstysIeuProGlnLeuAsnAlaAlaAglLeuGly	140		
Qy			

Db 774 AAAGACTCTGTTTATACAAATGAGAAACCTTCAGAAATCTCTGCCCCCAATCGAGCATCT 533

[illegible]

161	GluLeuGlnGlnSerArgLysMetPhe	AsrMetTyrAlaLeuLys---	ThrLeuGly	179
Qy				
885	ATGATCATCAAAACCCGCGAA	-----	TACGGCTTCAACACGACGACAC	929
Dh				
26	LysLeuLeuSerAlaGluValGluLysMetLeuValGln	YsGlyAlaProAsnGluGly	45	
Qy				
13056	AAATTCGACAGCTTCTACGTAAGAAACGAAACAA	-----	CTTCTCGGCGATAC	138
Dh				

Cy		18C GlyValValAsnThr i84 	
Cy		18Q GlyValValAsnThr i84 	
Cy	46	IleGluValValPheCysLeuLeuLeuTyAlaLeuAlaAargThrTrpSerProLys 65         :    :	
Cy	10000	HtttLgtGGAGCGCTTGTATGTTTGTTTTGGCTTCATTGTTTGTGGCTGAGAAG	

Db	930	GGATGTTGTGGACA	944	
				138
				139
				140
				141
				142
				143
				144
				145
				146
				147
				148
				149
				150
				151
				152
				153
				154
				155
				156
				157
				158
				159
				160
				161
				162
				163
				164
				165
				166
				167
				168
				169
				170
				171
				172
				173
				174
				175
				176
				177
				178
				179
				180
				181
				182
				183
				184
				185
				186
				187
				188
				189
				190
				191
				192
				193
				194
				195
				196
				197
				198
				199
				200
				201
				202
				203
				204
				205
				206
				207
				208
				209
				210
				211
				212
				213
				214
				215
				216
				217
				218
				219
				220
				221
				222
				223
				224
				225
				226
				227
				228
				229
				230
				231
				232
				233
				234
				235
				236
				237
				238
				239
				240
				241
				242
				243
				244
				245
				246
				247
				248
				249

AAF86431  
ID AAF86431 standard; DNA; 349980 bp.  
XX  
138125 -----AAAGATGGGACGTTGCTATAGCAAGGGCTTTGGATACAGAAATGTTGAG 138

	Qy	84	84
AC			
AAF86431;			
XX			
DT	29-OCT-2001	138176	138176
	(first entry)	GCAAGGCTACCATCTACCCCGAGACAATATACCGAATTGGTTCAATAACCAAGAGCTTT	138176

XX		
DE	Pyrococcus abyssi genomic fragment #1.	
XX		
XX		

---

Cy	85	--ValIValThrGluGlyAspLeuLysTyrValIleuAspGly97 :::   ::   ::
----	----	--

```

Ddb 138236 ACAGCCTTGGCAATGTAATAGTGGAGAGGCGGNTTAAGT-----TTGGATGAC 138289
Qy 98 CysAlaProLeuThrArgPheThrAsn---LysLeuArgThrPheGlyArgThrPheThr 116
Ddb 138290 -----CCGTTGAGAAATTCGTTAATATAAGCTTAGACCATTCGGAGAACCAAGTGACC 138343
Qy 117 G'uaLaTyValAspPheCys::ealaTyLysHisLysLeuProGlnLeuAsnAlaala 136
Ddb 138344 -----GTTCAACACCTGTAAACGATTCCTCGGGATTCACATCTTAGGATA-GCC 138394
Qy 137 AlaGluLeu-----GlyLeuProAlaGluAspSerTyLeu 148
Ddb 138395 GAGGCTTCATAGACGGAATGCTCGGTGGGATAACTGGTTG 138436

RESULT 9
AAT68648
ID AAT68648 standard; DNA; 4467 BP.
XX
AC AAT68648;
Df 19-AUG-1997 (first entry)
XX
DE PSRQ800 fragment including coding sequence for AbiE.
XX
KW AbiE; phage abortive infection protein; phage resistance; pSRQ800;
KW lactic acid bacterium; ss.
XX
OS Lactococcus lactis subsp. lactis isolate W1.
FF
FH Key Location/Qualifiers
FT -35_signal 1203..1207
FT /*tag= a
FT -10_signal 1225..1230
FT /*tag= b
FT RBS 1265..1271
FT /*tag= c
FT CDS 1279..3078
FT /*tag= d
XX
W0972C917-A2.
XX
X2 JUN-1997.
XX
20-NOV-1996; 96MO-IB01385.
XX
01-DEC-1995; 95US-0565907.
XX
(UNIL) QUEST INT BV.
XX
Holler BJ, Kondo JK, Moineau S, Vandenbergh PA;
Vedamuthu ER;
WPI: 1997-3:9765/29.
P-PSDB; AAU17788.
XX
Isolated DNA encoding the AbiE protein of Lactococcus - for
protecting strains used in production of fermented dairy products
XX
Claim 4; Page 29-32; 49pp; English.
XX
The 4.5 kb EcoRI fragment (AAT68648) of plasmid pSRQ800 includes
a coding sequence for AbiE (AAU17788), a protein which, in
Lactococcus lactis (L1), increases resistance to phages by
aborting infection. pSRQ800 was isolated through examination of
the phage resistance mechanism of L1 subsp. lactis W1. The total
plasmid DNA of W1 was co-electroporated with shuttle vector pSA3
into phage-sensitive, plasmid-free L. lactis LM0230. Phage-
resistant transformants were isolated and found to contain pSRQ800.
The isolated DNA can be used to impart phage resistance to
bacteria, esp. L1., for use in the prodn. of fermented dairy
products.
XX
Sequence 4467 BP; 1583 A; 668 C; 643 G; 1573 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 7.09 Length: 4467
Score: 85.00 Matches: 36
Percent Similarity: 41.89% Conservative: 26
Best Local Similarity: 24.32% Mismatches: 46
Query Match: 8.58% Indels: 40
DB: 16 Gaps: 6

US-09-613-486-15 (1-198) x AAT68648 (1-4467)
Qy 2 GluLeuMetSerAspSerAsnLeu-----SerAsnLeuValIleThrAsp--- 16
Ddb 2059 GAATTTTAAATGAATTTAATCTAATCTCTCGAGAAATAACTTAATTAATGATAT 2118
Qy 16 ----- 16
Ddb 2119 AAAACGAAAGTTGACAAATTCGCCGTTTGTGTATAATCGAGTAAATCGGATATTTTCT 2178
Qy 17 -----AlaSerSerLeuSerGlyValAspLys-----LysLeuSerAla 30
Ddb 2179 TTTTGTGAAATATTAATCTCACTAATTCACACGACAGTGGATTAAAGAAATAAGCAAT 2238
Qy 31 GluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGly---IleGluValVal 49
Ddb 2239 TTTATAGATTATTGCTGAATGAAGACATTTAGGCAATAGGAGCTATAAAATGTATT 2298
Qy 50 PheGlyLeuLeuLeuTyrrAlaLeuAlaAlaArgThrSerProLysValGlnArgAla 69
Ddb 2299 TTCCCAAGTTATA-----ACAAATACATTCGAAACAAAAAAGTA 2337
Qy 70 AspSerAspValIlePheSerSerPheGlyGluArgAsnValValValThrGluGly 89
Ddb 2338 GATCTATAAAATATA--GACATATCTTTTCGAAAGAAACATGGTTACCAATTTTAAAT 2394
Qy 50 AspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuArg 109
Ddb 2395 GTTTTCGAAAAATATTAGATTATCATTAAGAAATCAAGATTAACTTAAGTTTTC 2454
Qy 110 ThrPheGlyArgThrPheThrGlu 117
Ddb 2455 ACTTCTTTGAAAAATATTAAATGAA 2478

RESULT 10
AAA9:259
ID AAA91259 standard; DNA; 4766 BP.
XX
AC AAA91259;
Df 08-MAY-2001 (first entry)
XX
DE GLRAV-5 nuclease sequence.
XX
KW GLRAV-5; grapevine leafroll virus; GLRAV infection; GLRAV coat protein;
KW GLRAV HSP70 homologue protein; viral gene mapping;
KW plant disease resistance; ds.
XX
OS Grapevine leafroll virus.
XX
FH Key Location/Qualifiers
FT CDS 1..159;
FT /*tag= a
FT /product= "HSP70 homologue protein"
FT /partial
FT /note= "No start codon given; Specifically claimed region"
FT CDS 1574..2989
FT /*tag= b
FT /product= "ORF 2 protein"
FT /note= "Specifically claimed region"
FT CDS 3285..4094
FT /*tag= c
FT /product= "coat protein"

```

CDS /note= "Specifically claimed region"  
 4128..475:  
 /tag= d  
 /product= "duplicate coat protein"  
 FT /note= "Specifically claimed region"  
 FT  
 XX  
 WQ200105957-A2.

PN  
 XX  
 PD 25-JAN-2001.

XX 19-JUL-2000; 2000WO-US19708;

XX 19-JUL-1999; 99US-0144453.

PA (AGRI-) AGRITOPIC INC.

XX Good XC, Morris J;

XX WPI; 2001-147339/15.

DR P-PSDB; AAY97688, AAY97689, AAY97690, AAY97691.

XX Novel grapevine leafroll virus polynucleotide useful as diagnostic and  
 PT probe, for viral gene mapping and for induced plant disease resistance  
 PT

XX Claim 1; Fig 1; 60pp; English.

XX This sequence represents a grapevine leafroll virus (GLRaV-5) DNA  
 CC sequence of the invention. The DNA sequence can be used in an expression  
 CC construct. The construct is useful for providing resistance to GLRaV  
 CC infection in a recombinant plant cell by transforming the plant cell with  
 CC it, where transcription of the polynucleotide sequence interferes with a  
 CC normal viral function such as movement, encapsidation or replication of  
 CC viral RNA. The polynucleotide sequence is expressed as an antisense.  
 CC sequence and encodes a GLRaV coat protein, preferably a defective GLRaV  
 CC coat protein or a GLRaV HSP70 homologue protein. The GLRaV-5 DNA is  
 CC useful for the synthesis of GLRaV, as diagnostics and probes, for viral  
 CC gene mapping and for induced plant disease resistance. It is also useful  
 CC to detect and quantitate expression of GLRaV in plant tissue prior to use  
 CC in vegetative propagation, by detecting the presence of GLRaV RNA.

XX Sequence 4766 BP; 1329 A; 890 C; 1231 G; 1316 T; 0 other;

Alignment Scores:

Pred. No.:	9-03	Length:	4766
Score:	84.50	Matches:	32
Percent Similarity:	40.46%	Conservative:	21
Best Local Similarity:	24.43%	Mismatches:	57
Query Match:	8.53%	Indels:	21
DB:	22	Gaps:	4

US-09-613-486-15 (1-198); x AAA91259 (1-4766).

QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValillePheSerAsnSerPheGly 80  
 Db TCAACGTCCTCCCTAAGGTGTCACAGTCTCCACATCGGACTATA---ACCGTAAAGTAGATG 3719

QY 8: GLuArgAsnValValThrGluGlyAspLeuLysLysValleuAspGlyCysAlaPro 100  
 Db GGTAAAGGAAGTACAGTTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATG 3779

QY 10: LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120  
 Db TCTTCGGTTACGAAACACCTCTAGACATTTGGAGAGCATTCACGGCAGCGTAGTG 3839

QY 121 AspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
 Db CAAGGGATATCTCCAGGCAAACTGGAGGTG---AACACGAAGATCTGCGCTTCTCAGCGT 3896

QY 141 IleProAlaGlu---TCTTACTCGCCAGACTGTCTGCATGTTGATGCAAGACTGTTTC 144

Db GTGCCGCCCTAACTACTACTTCTTACTCGCCAGACTGTCTGCATGTTGATGCAAGACTGTTTC 3956

QY 145 -----AspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeu 162  
 Db 3957 GGTATGACGCTTCACCTGGCAGCTGAG---TTAGCGAAGATGGTTGGCTATCAATAAACCT 4013  
 QY 163 GlnGlnSerArgLysMetPheAlaSerMetTyr 173  
 Db 4014 TCCATATAGCAACAGAGCTACACACAATCTGTAC 4046

RESULT 11

ABA92787/C  
 ID ABA92787 standard; DNA; 640681 BP.

XX ABA92787;

XX 27-MAR-2002 (first entry)

XX Buchnera sp. genomic DNA SEQ ID NO:1.

XX Buchnera; cockroach-symbiotic bacterium; cockroach extermination;  
 XX circular; ds.

XX Buchnera sp.

XX JP2001292771-A.

XX 23-OCT-2001.

XX 07-APR-2000; 2000JP-0107160.

XX 07-APR-2000; 2000JP-0107160.

XX (RIKA) RIKAGAKU KENKYUSHO.

XX WPI; 2002-126043/17.

XX A genomic DNA of cockroach-symbiotic bacterium -

XX Claim 1; Page 16-230; 237pp; Japanese.

XX The present invention describes a gene (I) derived from Buchnera sp.  
 CC containing the DNA (a) or (b), (a) has a fully defined base pair  
 CC sequence selected from a table of sequences found in the Buchnera sp.  
 CC genomic DNA of ABA92787 given in the specification or is a DNA selected  
 CC from complementary DNA sequences, and (b) is a DNA which hybridises with  
 CC the DNA (a) and encodes a protein. Also described are: (1) a recombinant  
 CC vector (12) containing (1); (2) a transformant (11) containing (1);  
 CC (3) a genomic DNA of Buchnera sp. containing the sequence given in  
 CC ABA92787; (4) a plasmid derived from Buchnera sp. containing DNA (c) or  
 CC (d), (c) is a DNA containing a fully defined sequence given in ABA92785  
 CC or ABA92789 and (d) is a plasmid which hybridises with a DNA; and (5) a  
 CC method for the preparation of a protein in which (1) is cultured and  
 CC the expression protein of the objective protein is collected from the  
 CC resultant culture. The DNA is useful for developing agricultural  
 CC chemicals for exterminating cockroaches. The present sequence represents  
 CC the specifically claimed Buchnera sp. genomic DNA sequence, from the  
 CC present invention.

XX Sequence 640681 BP; 237522 A; 83922 C; 84757 G; 234580 T; 0 other;

Alignment Scores:

Pred. No.:	1.5e+04	Length:	640681
Score:	83.50	Matches:	32
Percent Similarity:	37.18%	Conservative:	26
Best Local Similarity:	20.51%	Mismatches:	53
Query Match:	8.43%	Indels:	45
DB:	24	Gaps:	5

US-09-613-486-15 (1-198) x ABA92787 (1-640681)

QY 65 LysValGlnArgAlaAspSerAspValillePheSerAsnSerPheGlyGluArgAsnVal 84

Db 435915 AAAAATCAACGTGGATCATTAAGCATTTAGCAAGAGTACTCCAGGTCAAAAGAAATATT 435856





fragment (AAB14800):"

```

XX JP3072345-B..
XX
XX 31-JUL-2000.
XX
XX 31-MAR-1999; 99CF-0694004.
XX
XX 31-MAR-1999; 99CF-0694004.
XX
XX (NORQ ) NORINSUSANSHO KACHIKU E-SEI SHIKENCOCCHO.
XX (HGET ) HIGETA SHOUYU KK.
XX (FUJ-) FUJITA GAKUEN.
XX
XX WPI: 2000-551432/51.
XX P-PSDB; AAB14800.
XX
XX Novel recombinant protective polypeptide antigen useful as a vaccine
XX for protecting livestock against Erysipelothrix rhusiopathiae
XX infection -
XX
XX Example 1; Fig 2; 23pp; Japanese.
XX
XX The invention relates to a 46.5 kD immunogenic fragment (AAB14800) of
XX the protective polypeptide antigen (PPA) of the bacterium Erysipelothrix
XX rhusiopathiae, and to DNA encoding it (AAV72313). This bacterium infects
XX livestock, particularly pigs (swine erysipelas), and is also able to
XX cause disease in humans. The recombinantly produced PPA fragment may be
XX used as a subunit vaccine for stimulating an immune response in animals
XX against Erysipelothrix rhusiopathiae. As the vaccine can be produced
XX recombinantly, it can be produced on a large scale. The vaccine is
XX suitable for transcutaneous administration. The present sequence represents
XX the coding sequence of DNA encoding the full-length PPA.
XX
XX Sequence 1881 BP; 719 A; 291 C; 384 G; 487 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 5.65 Length: 1881
XX Score: 81.50 Matches: 33
XX Percent Similarity: 40.60% Conservative: 21
XX Best Local Similarity: 24.81% Mismatches: 50
XX Query Match: 8.22% Indels: 29
XX DB: 21 Gaps: 5
XX
XX US-09-613-486-15 (1-198) x AAV72313 (1-1881)
XX
XX QY 60 ArgThrThrSerProLysValGlnArgAspSerAspValIlePheSerAsnSerPhe 79
XX
XX Db 445 AGATTACCGATTCTGAACTGTGATGAGCATACGTTGATTTACTCGTGAATTACGAGGTG 504
XX
XX QY 80 GlyGluArgAsnValValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAa 99
XX
XX Db 505 AAACACCGTATTTAGTAAATATAGAGTAAAGTTAA-----GGTAGAGT 552
XX
XX QY 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
XX
XX Db 553 CCCTTAGACGATTATTAGTCTCTCAAGACAGAGAATTCGTAGTGAATGAATTGCT 612
XX
XX QY 117 -----GluAlaIyValAspPheCysIleala----- 125
XX
XX Db 613 GCAGAGTAATATTATTACCTGAAGCGCATGAGGATTTCTTAGTTTCAGATTCAACGAG 672
XX
XX QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
XX
XX Db 673 TATAATCACAACTAAATAATATCAACTTTGCTTTGGGTCTAGGGTCACGGAGTTATT 732
XX
XX QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSer 165
XX
XX Db 733 GACTAT-----AACCGGCTCGAAATATGATGGAAAAA 765
XX
XX QY 166 -----ArgLysMetPheAlaSerMetTyrAlaLeuLys 176
XX
XX Db 766 GAATTCATCCACTGTATTGAACTTTATGCTATCGG 804

```

```

RESULT 14
AAAS0205
ID AAAS0205 standard; DNA; 2040 BP.
XX
XX AAAS0205;
XX
XX C7-NCV-2000 (first entry)
XX
XX Erysipelothrix rhusiopathiae erysipelas protective antigen gene.
XX
XX Erysipelas protective antigen; Epa; SpaA.1; DNA vaccine; infection;
XX immuno-protective epitope; ds.
XX
XX Erysipelothrix rhusiopathiae.
XX
XX Key Location/Qualifiers
XX -35_signal 41..46
XX /*tag= a
XX -10_signal 65..70
XX /*tag= b
XX RBS 86..92
XX /*tag= c
XX CDS 100..1980
XX /*tag= d
XX sig_peptide 100..186
XX /*tag= e
XX mat_peptide 187..1977
XX /*tag= f
XX
XX W0200047744-A1.
XX
XX 17-AUG-2000.
XX
XX 10-FEB-2000; 2000MO-US03789.
XX
XX 10-FEB-1999; 99US-0119389.
XX (UYRQ ) UNIV ROCKEFELLER.
XX
XX Fischetti VA, Shimoji Y;
XX
XX WPI: 2000-524541/47.
XX
XX P-PSDB; AAY95782.
XX
XX Vaccines for protecting turkeys and pigs against Erysipelothrix
XX rhusiopathiae infections comprising a polypeptide sequence from the
XX N-terminal region of an erysipelas protective antigen -
XX
XX Example; Fig 2; 61pp; English.
XX
XX The present sequence is that of the Epa gene of strain Fujisawa of
XX Erysipelothrix rhusiopathiae, the causative agent of erysipelas in
XX animals and erysipeloid in humans. The gene encodes a protective
XX antigen (see AAY95782), termed erysipelas protective antigen (Epa or
XX SpaA.1). It was isolated from a genomic DNA library of strain
XX Fujisawa by immunoscreening using E. rhusiopathiae convalescent
XX pig serum. The N-terminal portion of the Epa protein, especially a
XX polypeptide comprising residues 12-195 of the protein, was
XX identified as a vaccine antigen. The N-terminal polypeptide was
XX shown to protect mice and pigs from a lethal challenge with E.
XX rhusiopathiae. Vaccines containing immunogenic polypeptides
XX of E. rhusiopathiae, where the immunogenic polypeptide comprises
XX an immuno-protective epitope from the N-terminal region of Epa,
XX are claimed. A claimed method for protecting an animal,
XX especially a turkey or pig, from infection by E. rhusiopathiae
XX involves administering the vaccine, or an expression vector
XX comprising a nucleic acid encoding the N-terminal region or
XX full-length Epa protein.
XX
XX Sequence 2040 BP; 776 A; 293 C; 418 G; 553 T; 0 other;
XX
XX Alignment Scores:

```

```
Pred. No.: 6.36 Length: 2040
Score: 81.50 Matches: 33
Percent Similarity: 40.60% Conservative: 21
Best Local Similarity: 24.81% Mismatches: 50
Query Match: 8.22% Indels: 29
DB: 21 Gaps: 5

US-09-613-486-15 (1-198) x AAA50205 (1-2040):
Qy 60 ArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe 79
Db 544 AGATTACGATTCCTGAACATGATGAAGCATACGTTGATTACTCGTGAATTACGAGGTG 603
Qy 80 GlyGluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAla 99
Db 604 AACACCGGTATTTAGTAAATAATGAAGGTAAGTTAAA-----GGTAGAGCT 651
Qy 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
Db 652 CCCTTAGAGCATTTATAGTTCCTCTAAGAGATAGAAATTCGTAGTATGAATGAATGCT 711
Qy 117 -----GluAlaTyrValAspPheCysIleAla----- 125
Db 712 GCAGAGTAATATATTACCTGAAGCGCATGAGGATTCCTAGTTTCAGATTCAACGGAG 771
Qy 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 772 TATAATGACAACTAAATAATATCACTTTGCTTTGGGCTAGGGTCACGAGTTTAT 831
Qy 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 832 GACTAT-----AACCGGCTCGAAATAATGATGGAAAAA 864
Qy 166 -----ArgLysMetPheAlaSerMetTyrAlaLeuLys 176
Db 865 GAACCTTCATCCACTGTACTCTGAACTTTATGCTATGCGG 903

RESULT 15
AAA72317
ID AAA72317 standard; DNA; 2814 bp.
XX
XX AAA72317;
XX
XX 11-DEC-2000 (first entry)
XX
XX Erysipelothrix rhusiopathiae full-length PPA DNA.
XX
XX Protective polypeptide antigen; PPA; Fujisawa strain; swine erysipelas;
KW recombinant production; immunogenic; subunit vaccine;
KW transnucous administration; ds.
XX
XX Erysipelothrix rhusiopathiae.
XX Key Location/Qualifiers
XX CDS 400..2280
XX /*tag= a
XX /product= "Erysipelothrix rhusiopathiae full-length PPA"
XX /product= 487..1692
XX /*tag= b
XX /product= "Erysipelothrix rhusiopathiae 46.5 KD PPA
XX fragment (AAB14800)."
XX
XX JP3072345-Bi.
XX
XX 31-JUL-2000.
XX
XX 31-MAR-1999; 99JP-0034004.
XX
XX 31-MAR-1999; 99CP-0094004.
XX
XX (NORQ ) NORINSUSANSHO KACHIKU EISEI SHIKENJOCHO.
XX (HGET ) HIGETA SHOYU KK.
XX (FUJI-) FUJITA GAKUEN.
```

```
XX WPI; 2000-SS1432/51.
DR P-PSD3; AAB14800.
XX
XX Novel recombinant protective polypeptide antigen useful as a vaccine
PT for protecting livestock against Erysipelothrix rhusiopathiae
PT infection
XX
XX Example 1; Fig 1A-1B; 23pp; Japanese.
XX
XX The invention relates to a 46.5 kD immunogenic fragment (AAB-4800) of
CC the protective polypeptide antigen (PPA) of the bacterium Erysipelothrix
CC rhusiopathiae, and to DNA encoding it (AAA72313). This bacterium infects
CC livestock, particularly pigs (swine erysipelas), and is also able to
CC cause disease in humans. The recombinantly produced PPA fragment may be
CC used as a subunit vaccine for stimulating an immune response in animals
CC against Erysipelothrix rhusiopathiae. As the vaccine can be produced
CC recombinantly, it can be produced on a large scale. The vaccine is
CC suitable for transnucous administration. The present sequence represents
CC DNA encoding the full-length PPA.
XX
XX Sequence 2814 BP; 1048 A; 409 C; 558 G; 799 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 10.1 Length: 2814
Score: 81.50 Matches: 33
Percent Similarity: 40.60% Conservative: 21
Best Local Similarity: 24.81% Mismatches: 50
Query Match: 8.22% Indels: 29
DB: 21 Gaps: 5

US-09-613-486-15 (1-198) x AAA72317 (1-2814):
Qy 60 ArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe 79
Db 844 AGATTACGATTCCTGAACATGATGAAGCATACGTTGATTACTCGTGAATTACGAGGTG 903
Qy 80 GlyGluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAla 99
Db 904 AACACCGGTATTTAGTAAATAATGAAGGTAAGTTAAA-----GGTAGAGCT 951
Qy 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
Db 952 CCCTTAGAGCATTTATAGTTCCTCTAAGAGATAGAAATTCGTAGTATGAATGAATGCT 1011
Qy 117 -----GluAlaTyrValAspPheCysIleAla----- 125
Db 1012 GCAGAGTAATATATTACCTGAAGCGCATGAGGATTCCTAGTTTCAGATTCAACGGAG 1071
Qy 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 1072 TATAATGACAACTAAATAATATCACTTTGCTTTGGGCTAGGGTCACGAGTTTAT 1131
Qy 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
Db 1132 GACTAT-----AACCGGCTCGAAATAATGATGGAAAAA 1164
Qy 166 -----ArgLysMetPheAlaSerMetTyrAlaLeuLys 176
Db 1165 GAACCTTCATCCACTGTACTCTGAACTTTATGCTATGCGG 1203
```

Search completed: November 8, 2003, 01:41:31  
Job time : 383 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 7, 2003, 23:44:24 ; Search time 63 Seconds  
(without alignments)  
1387.204 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991  
Sequence: 1 XELMSDLSNLVITDASSL.....GGVNTVPVSLRQLGRVYM 198

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame\_p2n,model:-DBVxih  
-Q/cgn2\_1/USP10.spool/US9613486/runat\_c7112003\_120411\_27115/app.query.fasta\_1.391  
-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=C.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BL0SUM62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OJFMT=ptc -NORV=ext -HEAPSIZE=500 -MINLEN=C -MAXLEN=200000000  
-USER=US09613486 @CNG 1.1.56 @runat\_07112003\_120411\_27115 -NCPU=6 -CFPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=C -WAIT -DSPBUCK=100 -LONGLOG  
-DBV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : issued Patents NA:

- 1: /cgn2\_6/prodata/2/ina/5A.COMB.seq:
- 2: /cgn2\_6/prodata/2/ina/5B.COMB.seq:
- 3: /cgn2\_6/prodata/2/ina/6A.COMB.seq:
- 4: /cgn2\_6/prodata/2/ina/6B.COMB.seq:
- 5: /cgn2\_6/prodata/2/ina/PCUS.COMB.seq:
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	991	100.0	597	3	US-09-080-983-14
2	991	100.0	15500	3	US-09-080-983-1
3	86	8.7	2034	4	US-09-252-991A-1531
4	85	8.6	4467	1	US-08-565-907A-1
5	85	8.6	4467	2	US-08-910-551B-1
6	85	8.6	4467	2	US-08-909-425A-1
7	81	8.2	1743	4	US-09-252-991A-13408
8	81	8.2	2139	4	US-09-252-991A-13503
9	81	8.2	2433	4	US-09-252-991A-13950
10	79	8.0	1434	2	US-08-770-544-15
11	79	8.0	1434	4	US-09-579-259-15
12	79	8.0	536165	4	US-09-214-808-1

13	77.5	7.8	1233	4	US-09-252-991A-3735	Sequence 3735, Ap
14	77.5	7.8	1491	4	US-09-252-991A-3635	Sequence 3636, Ap
15	77.5	7.8	4884	4	US-09-252-991A-3824	Sequence 3824, Ap
16	77	7.8	1776	4	US-09-252-991A-15987	Sequence 15987, A
17	77	7.8	1794	4	US-09-252-991A-16086	Sequence 16086, A
18	77	7.8	1878	4	US-09-252-991A-16516	Sequence 16516, A
19	76.5	7.7	7676	1	US-08-451-777A-7	Sequence 7, Appl
20	76.5	7.7	7676	2	US-08-451-778A-7	Sequence 7, Appl
21	76.5	7.7	7676	2	US-08-998-208-7	Sequence 7, Appl
22	76.5	7.7	7676	5	PCR-US95-06743-7	Sequence 7, Appl
23	76	7.7	672	3	US-09-080-983-12	Sequence 12, Appl
24	75	7.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
25	74.5	7.5	1764	4	US-09-107-532A-2751	Sequence 2751, Ap
26	74	7.5	1173	3	US-09-285-601-1	Sequence 1, Appl
27	74	7.5	3515	1	US-08-596-985-1	Sequence 1, Appl
28	74	7.5	6709	3	US-09-285-601-3	Sequence 3, Appl
29	73.5	7.4	2504	1	US-08-484-105-15	Sequence 15, Appl
30	73.5	7.4	2504	1	US-08-484-106-15	Sequence 15, Appl
31	73.5	7.4	2680	3	US-08-949-155-50	Sequence 50, Appl
32	73.5	7.4	2680	4	US-09-819-964-50	Sequence 50, Appl
33	73.5	7.4	3260	4	US-09-221-017B-980	Sequence 980, App
34	73.5	7.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl
35	73.5	7.4	4411529	3	US-09-103-840A-1	Sequence 1, Appl
36	73	7.4	49136	3	US-09-422-869-1	Sequence 1, Appl
37	73	7.4	59065	4	US-09-813-817-3	Sequence 3, Appl
38	73	7.4	59065	4	US-09-978-197-3	Sequence 3, Appl
39	72.5	7.3	3355	4	US-09-221-017B-412	Sequence 412, App
40	72.5	7.3	7172	4	US-08-961-527-120	Sequence 120, App
41	72	7.3	2746	2	US-08-576-165-3	Sequence 3, App
42	72	7.3	7425	4	US-09-453-702B-212	Sequence 212, App
43	71.5	7.2	334	3	US-09-060-756-261	Sequence 261, App
44	71.5	7.2	334	4	US-09-670-314-261	Sequence 261, App
45	71.5	7.2	1386	4	US-09-252-991A-15531	Sequence 15531, A

ALIGNMENTS

RESULT 1

Sequence 14, Application US/09380983

Patent No. 6197948

GENERAL INFORMATION:

APPLICANT: Zhu, Hai-Ying

APPLICANT: Ling, Kai-Shu

APPLICANT: Gopalves, Dennis

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS

TITLE OF INVENTION: AND THEIR USES

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER:

APPLICATION NUMBER: US/09/080,983

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA: US 60/047,194

APPLICATION NUMBER: US 60/047,194

FILING DATE: 20-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/1631

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 597 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

US-09-080-983-14

Alignment Scores:  
 Pred. No.: 1,016-127 Length: 537  
 Score: 991.00 Matches: 198  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-613-486-15 (1-198) x US-09-080-983-14 (1-597)

```

QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db 1 ATGGAGTGTGATGCGCAGCAGCACTTACCACTCTGCTGATTAACCGCGCTCTAGTCTA 60
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
Db 61 AATGGTGTGCGACAGAAGCTTTTATCTGCTGAAGTTGAAAAAATGTTGTCAGAAAGG 120
QY 41 AApCAsnGluGlyIleGluValValPheGlyLeuLeuLeuValAlaAlaArg 60
Db 121 GCTCCTAACGAGGGTATAGAAGTGTGTTGCGTCTACTCCTTACGCACCTCGCGCAAG 180
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 90
Db 181 ACCACGTCCTCCTAGAGTTTCAGCGCGCAGATTCAGACGTTATATTTCAAAATGTTT 240
QY 81 GluArgAsnValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100
Db 241 GAGAGGAATGTGTAGTAAACAGAGGGTGACCTTAAGAGAGTACTCGACGGGTGTCGCG 300
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaVal 120
Db 301 CTCACCTAGGTTCACTANTAACTTGAACGTTTCGGTCTGCTCTTCTCCTGAGGCTTAC 360
QY 121 AspPheCysIleAlaTyrrLysHisLysLeuProGlnLeuAsnAlaAlaAlaGlu 140
Db 361 GACTTTTGTATCGGTATTAAGCACAATTAACCCCACTCAACCGCGCGCGAATTGGGG 420
QY 141 IleProAlaGluAspSerTyrrLeuAlaAlaAspPheLeuGlyThrCysProLysLeu 160
Db 421 ATCCAGCTGAAGATTCGTACTTACCTGACAGATTTTCTGGGTACTTGGCCGAAGCTCT 480
QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrrAlaLeuLysThrGluGly 180
Db 481 GAATTACAGCAAGTAGGAGATGTTTCGGAGTATGTACGCTCTTAAAACTGAAGTGG 540
QY 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgGluValMet 198
Db 541 GTGGTAATACACCAAGTACGCAATCTGCTGACAGTACGCTAGGTAGAGGAAGTATG 594
  
```

RESULT 2

US-09-080-983-1

Sequence 1, Application US/C9080983  
 Patent No. 6197948  
 GENERAL INFORMATION:  
 APPLICANT: Zhu, Hai-Ying  
 APPLICANT: Ling, Kai-Shu  
 APPLICANT: Gonsalves, Dennis  
 TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS  
 TITLE OF INVENTION: AND THEIR USES  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon, Hartgrave, Devans & Doyle LLP

STREET: Clinton Square, P.O. Box 1051  
 CITY: Rochester  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 14603  
 COMPUTER READABLE FORM:  
 MEDIA TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/080,933  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA: US 60/047,194  
 APPLICATION NUMBER: 20-MAY-1997  
 FILING DATE: 20-MAY-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Goldman, Michael L.  
 REGISTRATION NUMBER: 30,727  
 REFERENCE/DOCKET NUMBER: 19603/1631  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (716) 263-1304  
 TELEFAX: (716) 263-1600  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15500 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

US-09-080-983-1

Alignment Scores:  
 Pred. No.: 1,566-125 Length: 15500  
 Score: 991.00 Matches: 198  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-613-486-15 (1-198) x US-09-080-983-1 (1-15500)

```

QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
Db 13584 ATGGAGTGTGATGTCGACAGCACTTACCACTTGGTATTAACCGCGCTCTAGTCTA 13643
QY 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40
Db 13644 AATGGTGTGCGACAGAAGCTTTTATCTGCTGAAGTTGAAAAAATGTTGTCAGAAAGG 13703
QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrrAlaLeuAlaArg 60
Db 13704 GCTCCTAACGAGGGTATAGAAGTGTGTTGCGTCTACTCTTTACGCACCTCGCGGCAAG 13763
QY 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
Db 13764 ACCACGTCCTCCTAAGGTTTCAGCGCGCAGATTCAGACGTTATATTTTCAAAATAGTTTCGGA 13823
QY 81 GluArgAsnValValThrGluGlyAspLeuLysLysValIleAspGlyCysAlaPro 100
Db 13824 GAGAGGAATGTGGTAGTAACAGAGGCTGACCTTAAGAGAGTACTCGACGGGTGTGCGCT 13883
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaVal 120
Db 13884 CTCACCTAGGTTCACTAATAAACTTAGAACTTGGTCTGCTACTTCTCAGTACGCTTACGTT 13943
QY 121 AspPheCysIleAlaTyrrLysHisLysLeuProGlnLeuAsnAlaAlaAlaGluLeuGly 140
Db 13944 GACTTTTGTATCGGTATTAAGCACAATTAACCCCACTCAACCGCGCGCGAATTGGGG 14003
QY 141 IleProAlaGluAspSerTyrrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160
  
```

```
Db 14004 ATTCCAGCTGAGATTGCTGCTTAGCTGCGAATTTCTGGGTACTTCCCGGAGGCTCTCT 14063
Qy 161 GluLeuGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180
Db 14064 GAATTACAGCAAGTAGAAGAGTTTCGGGAGTATGTACGCTCTAAAACTGAGGTCGA 14123
Qy 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluValMet 198
Db 14124 GTGGTAATACACAGTGGAGCAATCTCGGTACGCTAGGTAGGAGGAAGTTATG 14177

RESULT 3
US-09-252-991A-1531/c
; Sequence 1531, Application US/03252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1531
; LENGTH: 2034
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1531

Alignment Scores:
Pred. No.: 0.1 Length: 2034
Score: 86.00 Matches: 35
Percent Similarity: 43.04% Conservative: 33
Best Local Similarity: 22.15% Mismatches: 76
Query Match: 6.88% Indels: 14
DB: 4 Gaps: 4

US-09-613-486-15 (i-198) x US-09-252-991A-1531 (1-2034)
Qy 24 AspLysLysLeuLeuSerAaGluValGluLysMetLeuValGlnLysGlyAaProAsn: 43
Db 864 GATCAGCGCTCCGGCTTCGCGGCTCAGCGGAGTCGCCGATGAGCGCGCCAGGA 865
Qy 44 GluGlyTleGluValValPheGlyLeuLeuTyrAlaLeuAlaAaArgThrThrSer 53
Db 804 GCGGGCATGAGACGATCGCGGGGACCTCGTCGCAACCATCGCGGTCCGCGGT 745
Qy 64 ProLysValGlnArgAaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsn 53
Db 744 GACCAGCTTGCCACGCGGAGCTGGCTTGTGAGAACACACCGCGGGACAGGCCAG 585
Qy 84 ValValValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArg 103
Db 684 GCCGGGTAGAAATCGGT-----GTCAAGGCTGT----- 655
Qy 104 PheThrAsnLysLeuAaThrPheGlyArgThrPheThrGluAlaTyrValAspPheCys 123
Db 654 ---CTCAAGGCGCTGAGCTCGACCGCCGAGCGGTTCAGCAGGTGTTCGCGTTC 598
Qy 124 IleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAa---AlaGluLeuGlyIlePro 142
Db 597 ATACAGGGCGGTTCGGCATTCACCGCATTCGTCGCCCGCTAGCCGAGGATAGCGCCC 538
Qy 143 AlaGluAspSerTyrLeuAaAlaAspPheLeuGlyThrCysProLysLeuSerGLeu 162
Db 537 GCCGGCTTGAATTCGTTGGCTTGGATCGCCGCCAAGTCGTGCTTCATGATGAC 478
Qy 163 GlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180
Db 477 CAGGATCGCGCCCGC-----GGGGTCCGCTTCGGATAGAACACAGCGCGCGC 430

US-08-565-907A-1
; Sequence 1, Application US/08565907A
; Patent No. 5814499
; GENERAL INFORMATION:
; APPLICANT: Sylvain Moineau, Barbara
; APPLICANT: J. Holler, Peter A. Vandenberg,
; APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.
; APPLICANT: Kondo
; TITLE OF INVENTION: DNA Encoding Phage
; TITLE OF INVENTION: Abortive Infection Protein
; TITLE OF INVENTION: From Lactococcus
; TITLE OF INVENTION: Lactis, and Method of Use Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 5.25 inch,
; MEDIUM TYPE: 360 Kb storage
; COMPUTER: Acer
; OPERATING SYSTEM: MS-DOS (version 4)
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/565,907A
; FILING DATE: December 1, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: Quest 4.1-152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: NO. 58144996
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4467
; TYPE: Nucleotide
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
; HYPOTHEICAL: No
; ANTI-SENSE: No N/A
; FRAGMENT TYPE: N/A
; ORIGINAL SOURCE:
; ORGANISM: Lactococcus lactis
; STRAIN:
; INDIVIDUAL ISOLATE: W1
; DEVELOPMENTAL STAGE: N/A
; HAPLOTYPE: N/A
; TISSUE TYPE: N/A
; CELL TYPE: bacterium
; CELL LINE: N/A
; ORGANELLAE: N/A
; IMMEDIATE SOURCE:
; LIBRARY: genomic
; CLONE: SMQ-20
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: phage abortive infection
; LOCATION: N/A
; IDENTIFICATION METHOD: sequencing
; OTHER INFORMATION: DNA encoding phage
```

OPERATING SYSTEM: MS-DCS (version 4)  
SOFTWARE: Wordperfect 5.1

SOFTWARE: Wordperfect 5.1

Db 2239 TTTATAGATTATGCTGTAATGAAGAACATTTAGGGAATTAAGGAGCTATAAATGCTATT 2298  
 QY 50 PheGlyLeuLeuLeuTyraLeuAlaAlaArgThrSerProLysValGlnArgala 69  
 Db 2299 TTCCCACTTATA-----ACAAATACATTTGAACACAAAAAAGTA 2337  
 QY 70 AspSerAspValIlePheSerAsnSerPheGlyGluArgAsnValValThrGluGly 89  
 Db 2338 GATACATAAAATATA---GACATATCTTTTGGAAAAGAAACATGTTACCAATTTTAA 2394  
 QY 90 AspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThrAsnLys-euArg 109  
 Db 2395 GTTTTCGAAAAAATATTAGATTTATCATTTAAAAGATTCAAGATTAACTAATTAAGTTTIG 2454  
 QY 110 ThrPheGlyArgThrPheThrGlu 117  
 Db 2455 ACCTTCTTTGAAAAATATTAAATGAA 2478

# RESULT 6

US-08-909-425A-1 ; Sequence 1, Application US/08909425A  
 ; Patent No. 5928688  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sylvain Moineau, Barbara  
 ; APPLICANT: J. Holler, Peter A. Vandenbergh,  
 ; APPLICANT: Ebenezer R. Vedamuthu, Jeffrey K.  
 ; APPLICANT: Kondo  
 ; TITLE OF INVENTION: DNA Encoding Phase  
 ; TITLE OF INVENTION: Abortive Infection Protein  
 ; TITLE OF INVENTION: From Lactococcus  
 ; NUMBER OF SEQUENCES: 1  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Ian C. McLeod  
 ; STREET: 2190 Commons Parkway  
 ; CITY: Okemos  
 ; STATE: Michigan  
 ; COUNTRY: USA  
 ; ZIP: 48864  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette 5.25 inch,  
 ; MEDIUM TYPE: 360 Kb storage  
 ; COMPUTER: Acer  
 ; OPERATING SYSTEM: MS-DOS (version 4)  
 ; SOFTWARE: Wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/909,425A  
 ; FILING DATE: August 11, 1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/565,907  
 ; FILING DATE: December 1, 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ian C. McLeod  
 ; REGISTRATION NUMBER: 20,931  
 ; REFERENCE/DOCKET NUMBER:  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (517) 347-4100  
 ; TELEFAX: (517) 347-4103  
 ; TELEX: No. 5928688e  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4467  
 ; TYPE: Nucleotide  
 ; STRANDEDNESS: Single  
 ; TOPOLOGY: Linear  
 ; MOLECULE TYPE:  
 ; DESCRIPTION: Genomic DNA  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FRAGMENT TYPE: N/A  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Lactococcus lactis

Alignment Scores:  
 Pred. No.: C.466 Length: 4467  
 Score: 85.00 Matches: 36  
 Percent Similarity: 41.89% Conservative: 26  
 Best Local Similarity: 24.32% Mismatches: 46  
 Query Match: 8.58% Indels: 40  
 DB: 2 Gaps: 6

US-09-613-486-15 (1-198) x US-08-909-425A-1 (1-4467)

QY 2 GluLeuMetSerAspSerAsnLeu-----SerAsnLeuValIleThrAsp--- 16  
 Db 2059 GAATTTTTAAATGAATTAATCTAACTCTGTCGAGAAATTAATTAATTAATGATTAAT 2118  
 QY 16 ----- 16  
 Db 2119 AAAACGAAAGTTGACAAATTTCCCGTTTGTGTATAAATCGAGTAAATCGGATATTTTCT 2178  
 QY 17 -----AlaSerSerLeuAsnGlyValAspLys-----LysLeuLeuSerAla 30  
 Db 2179 TTTTTCGAAAAATATTACTTCAACTAATTTCCACGACAGTGGATTAAAGAAATAGCAAT 2238  
 QY 31 GluValGluLysMetLeuValGlnLysGlyAlaProAsnGluGly---IleGluValVal 49  
 Db 2239 TTTATAGATTATTTGTGTGAATGAAGAACATTTAGGGAATTAAGGAGCTATAAATGTATT 2298  
 QY 50 PheGlyLeuLeuLeuTyraLeuAlaAlaArgThrSerProLysValG-nArga-a 69  
 Db 2299 TTCCCACTTATA-----ACAAATACATTTGAACACAAAAAAGTA 2337  
 QY 70 AspSerAspValIlePheSerAsnSerPheGlyGluArgAsnValValThrGluGly 89  
 Db 2338 GATACATAAAATATA---GACATATCTTTTCGAAAAGAAACATGTTACCAATTTTAA 2394  
 QY 90 AspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuArg 109  
 Db 2395 GTTTTCGAAAAAATATTAGATTTATCATTTAAAAGATTCAAGATTAACTAATTAAGTTTIG 2454  
 QY 110 ThrPheGlyArgThrPheThrGlu 117  
 Db 2455 ACCTTCTTTGAAAAATATTAAATGAA 2478

# RESULT 7

US-09-252-991A-13408  
 ; Sequence 13408, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Ruberfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A



```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13408
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13408

Alignment Scores:
Pred. No.: 0.389 Length: 1743
Score: 81.00 Matches: 30
Percent Similarity: 47.79% Conservative: 26
Best Local Similarity: 22.39% Mismatches: 50
Query Match: 8.17% Indels: 28
DB: 4 Gaps: 5

US-09-613-486-15 (1-198) x US-09-252-991A-13408 (1-1743)
QY 17 AlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeu 36
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 68 GCCAACACCATGAACGGTGTCTACCGCGAGGCGATCGCGAAGACCATCGACGGTGGAG 127
QY 37 ValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeuLeuTyrAla 56
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 128 GCGGAGAAA-----GAGGGGATCGCC-----GGCTGGTGGTGC 160
QY 57 LeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSer 76
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 161 -----ACCTCGCGAAGAGACCTTCTCGCGGCGGCGATCTC----- 199
QY 77 AsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLysValLeuAsp 96
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 200 -----AACGAGCTGATCAAGGTCACCAAGCGCGCGACCGCCCGGCTTCTACCCAG 247
QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 248 GGCATCTCGAAGCTCAAGGCGCAACTCGCGGCGCTGGAGACCTCGGCAAGCGGTGGTC 307
QY 117 GluAlaTyr-----ValAspPheCysIleAlaTyrLysHis 128
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 308 GCGCGCATCAACGGCGCTCGCTGGCGGCGGCTGGAGATCTGCTGGCTGCCACCCAC 367
QY 129 LysLeuProGlnLeuAsnAlaAlaAlaGluLeuGlyLeuPro 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 368 CGCATCGCCCTGGACATCCCGCGCTGGAGCTCGGCTGCCG 409

RESULT 8
US-09-252-991A-13503
; Sequence 13503, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13503
; LENGTH: 2199
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13503

Alignment Scores:
Pred. No.: 0.651 Length: 2199
Score: 81.00 Matches: 30
Percent Similarity: 41.79% Conservative: 26
Best Local Similarity: 22.39% Mismatches: 50
Query Match: 8.17% Indels: 28
DB: 4 Gaps: 5

US-09-613-486-15 (1-198) x US-09-252-991A-13503 (1-2199)
QY 17 AlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeu 36
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 GCCAACACCATGAACGGTGTCTACCGCGAGGCGATCGCGAAGACCATCGACGGTGGAG 189
QY 37 ValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeuLeuTyrAla 56
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 190 GCGGAGAAA-----GAGGGGATCGCC-----GGCTGGTGGTGC 222
QY 57 LeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSer 76
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 -----ACCTCGCGAAGAGACCTTCTCGCGGCGGCGATCTC----- 261
QY 77 AsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLysValLeuAsp 96
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 262 -----AACGAGCTGATCAAGGTCACCAAGCGCGCGACCGCCCGGCTTCTACCCAG 309
QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 310 GGCATCTCGAAGCTCAAGGCGCAACTCGCGGCGCTGGAGACCTCGGCAAGCGGTGGTC 369
QY 117 GluAlaTyr-----ValAspPheCysIleAlaTyrLysHis 128
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 370 GCGCGCATCAACGGCGCTCGCTGGCGGCGGCTGGAGATCTGCTGGCTGCCACCCAC 429
QY 129 LysLeuProGlnLeuAsnAlaAlaAlaGluLeuGlyLeuPro 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 430 CGCATCGCCCTGGACATCCCGCGCTGGAGCTCGGCTGCCG 471

RESULT 9
US-09-252-991A-13950/c
; Sequence 13950, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13950
; LENGTH: 2433
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13950

Alignment Scores:
Pred. No.: 0.651 Length: 2433
Score: 81.00 Matches: 30
Percent Similarity: 41.79% Conservative: 26
Best Local Similarity: 22.39% Mismatches: 50
Query Match: 8.17% Indels: 28
DB: 4 Gaps: 5

US-09-613-486-15 (1-198) x US-09-252-991A-13950 (1-2433)
QY 17 AlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeu 36
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2328 GCCAACACCATGAACGGTGTCTACCGCGAGGCGATCGCGAAGACCATCGACGGTGGAG 2269

```

QY 37 ValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyraAa 56  
 Db 2268 GCGGAGAA-----GAGGGATCGCC-----GGGTGGTGTG----- 2236  
 QY 57 LeuAaAlaArgThrSerProLysValGlnArgAlaAspSerAspValIlePheSer 76  
 Db 2235 -----ACCTCGCGAAGAAGACCTTCTTCGCGCGCGGATCTC----- 2197  
 QY 77 AsnSerPheGlyGlyLysAsnValValThrGluGlyAspLeuLysValLeuAsp 96  
 Db 2196 -----AAGAGCTGATCAAGGTTCACCAAGCGCGACCCCGCCCTTACCAG 2149  
 QY 97 GlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr 116  
 Db 2148 GGCATCTCGAATCAAGGGCAACTGCGCGCCCTGGAGACCTCGGCAAGCGGTGCTC 2089  
 QY 117 GluAlaTyR-----ValAspPheCysIleAlaTyRlyshis 128  
 Db 2088 GCGCGGATCAACGGCGCTGCGTGGCGCGCGCTGGAGATCTGCTGCGCTGCCACCAC 2029  
 QY 129 LysLeuProGlnLeuAsnAlaAlaGluLeuGlyIlePro 142  
 Db 2028 CCATCGCGCTGGACATCCCGCGCTGCGAGCTCGCGCTGCCG 1987

## RESULT 10

US-08-770-544-15

Sequence 15, Application US/08770544

Patent No. 5907085

GENERAL INFORMATION:

APPLICANT: Gonsalves, Dennis

APPLICANT: Ling, Kai-Shu

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND

NUMBER OF INVENTION: THEIR USES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans &amp; Doyle LLP

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DCS

SOFTWARE: Patent It. Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/770,544

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 6000908

FILING DATE: 21-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/621

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 1434 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-770-544-15

Alignment Scores:

Pred. No.: 0.543

Score: 79.00

Percent Similarity: 40.49%

Length: 1434

Matches: 43

Conservative: 40

Best Local Similarity: 20.98% Mismatches: 60  
 Query Match: 7.97% Indels: 62  
 Gaps: 2

US-09-613-486-15 (1-198) x US-08-770-544-15 (1-1434)

QY 12 LeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuLeuSerAlaGlu 31  
 Db 904 TTCCCATACGAAGCTCTACAGATAAC-----GCAAGA 939  
 QY 32 ValGluLysMetLeuValGlnLysGly-----AlaProAsnGly-----GlyIleGlu 47  
 Db 940 CTACGAGACTCGTCTTCGAAAGCGGGAGCAACACACGAGATATGGGAATATG 999  
 QY 48 ValValPheGlyLeuLeuLeuTyraAlaLeuAlaAlaArgThrThrSerProLysValGln 67  
 Db 1000 ATAGTGGCATGATACAACTTTTCGTACTCTCTCTACTGTAAAGAAATATAAGCGTCAAA 1059  
 QY 68 -----ArgAlaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsnVal 84  
 Db 1060 GACGGGTATAGGTGGAGACCGAATTA-----GGTCAAGAGAGATC 1101  
 QY 85 ValValThrGluGlyAspLeuLysValLeu-----AspGlyCysAlaPro 100  
 Db 1102 TACTTAAGTTATTTCGAAGTAAGGAAGCTATATTAGGAGGAAATACGTGCGTCTCCA 1161  
 QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyVal 120  
 Db 1162 -----ACCAACACTGTCGATCTCTCATGAGGTATTTTCTCACCACCTAT 1209  
 QY 121 AspPheCysIleAlaTyRlyshLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
 Db 1210 ACTCTACTTATACAGAAGAAATTCAG---CCAGCGTGTACTGCCCTAGTAAGCACGGC 1266  
 QY 141 IlePro-----AlaGluAspSerTyLeu 148  
 Db 1267 GTCCCGAAGAGTTCACTCGTACTGCTCGACTTCGACTCGATACACAGATATTAC 1326  
 QY 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMet 168  
 Db 1327 CCGCGGACGTGTG-----AAGGCTAACGCA 1353  
 QY 169 PheAlaSerMetTyraLeuLysThrGluGlyValValAsnThrProValSerAsn 188  
 Db 1354 ATGCTTGGCTATAGCATTAATCA-----GCTAAT 1386  
 QY 189 LeuArgGlnLeuGly 193  
 Db 1387 TTAAGGCGTAAAGGT 1401

## RESULT 11

US-09-579-259-15

Sequence 15, Application US/09579259

Patent No. 6558953

GENERAL INFORMATION:

APPLICANT: Gonsalves, Dennis

APPLICANT: Ling, Kai-Shu

TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS

PROTEINS AND THEIR USES

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans &amp; Doyle LLP

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent It. Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/579,259  
 FILING DATE: 25-May-2000  
 CLASSIFICATION: <UNKNOWN>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60509008  
 FILING DATE: 21-DEC-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Goldman, Michael L.  
 REGISTRATION NUMBER: 36,727  
 REFERENCE/DOCKET NUMBER: 19603/621  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (716) 263-1364  
 TELEFAX: (726) 263-1600  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 -LENGTH: 1434 base pairs  
 -TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
 09-579-259-15

Alignment Scores:		
Seq. No.:	0.543	Length:
Score:	79.00	Matches:
Percent Similarity:	40.4%	Conservative:
Percent Local Similarity:	20.9%	Mismatches:
Very Match:	7.97%	Indels:
:	4	Gaps:

1-09-613-486-15 (1-198) x US-09-579-259-15 (1-1434);

12	LeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGlu	31
904	TTGCGGATACGGAGGCTCTACAGATAAAC-----GCAAGA	939
32	ValGluLysMetLeuValGlnLysGly-----AlaProAsnGlu---	511
940	CTACGCACACTGTTCTTCGAAAGCGGGAGTCAACACACGAGATATGGGGAATATG	999
48	ValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLysValGln	67
1000	ATACTGGCCATGATACAACTTTTCGTACTCTACTCTACTGTAAGAATATAACGGTCAAA	1059
68	-----ArgAlaAspSerAspValLePheSerAsnSerPheGlyGluArgAsnVal	84
1060	GACGGGTATAGGGTGGACGGGAATTA-----GGTCAAAAGAGAGTC	1101
85	ValValThrGluLysAspLeuLysLysValLeu-----AspGlyCysAlaPro	100
1102	TACTTAAGTTATCGGAAGTAAAGGAGCTATATTAGAGGAAATACGGTGGCTGTCCA	1161
101	LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal	120
1162	-----ACCAACACTGTCCGATCCTTCATGAGGTATTTCGTCCACACCACTATT	1209
121	AspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly	140
1210	ACTCTACTATTAGAGAGAAAATTTCAG---CCAGCGTGTACTGCGCTAGTAAACACGGC	1266
141	IlePro-----AlaGluAspSerTyrLeu	148
1267	GTCCGGAAGAGGTCACTCCGTACTGTTTCGACTTCGACTCTGGATATACAGATTATAC	1326
149	AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMet	168
1327	CCGCGGACGGTGTG-----AAGGCTAACGCA	1353
169	PheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrProValSerAsn	188
1354	ATGGCTTCGGCTATAGCGAATAAATCA-----GGTAAT	1386

Qy	189	LeuArgGlnLeuGly	193
		::	
Db	1387	TAAAGGCGTAAAGCT	1401

RESULT 12

```

US-09-214-808-1
; Sequence 1, Application US/09214808A
; Patent No. 6475793
; GENERAL INFORMATION:
; APPLICANT: Rosenthal, Andre
; APPLICANT: Freiberg, Christoph
; APPLICANT: Perret, Xavier Philippe
; APPLICANT: Broughton, William John
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
; Patent No. 6475793
; TITLE OF INVENTION: Plasmid
; FILE REFERENCE: CARP0068
; CURRENT APPLICATION NUMBER: US/09/214,938A
; CURRENT FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/IB97/00950
; PRIOR FILING DATE: 1997-07-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO:
; LENGTH: 536165
; TYPE: DNA
; ORGANISM: Rhizobium
US-09-214-808-1

Alignment Scores:
Pred. No.: 5,2e+03
Score: 79.00
Percent Similarity: 38.27%
Best Local Similarity: 26.02%
Query Match: 7.97%
Gaps: 4
DE: 9

```

US-03-613-486-75 (1-198) X US-09-214-808-2 (1-536765)

Qy	1	MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerGerLeu	20
Db	314692	A TCGAAAAGCTTGGCGTTCTTAGTCTCCGGTCATGATTACCGACAAGCC-----	314742
Qy	21	AsnG:yValAspLysLysLeuSerAlaGiUValGlulYsMetLeuValGln:ySgLy	40
Db	314743	---GGAACCGCAAAAAAATCTGGCAGGGGTATTCATGAAGCTGCCAGAGCTTCGGCG	314799
Qy	41	-----Alap:GAsr: 43	
Db	314800	AGCAAGCTCATCAGCTCAGAAGACGCAAGTGGGAAATCGCTTTTGCTTCGCCCGAGC	314859
Qy	44	GluGIylleGluValvalPheGlyLeuLeuLeuTyRAlaLeuAlaalargThrTrsr	63
Db	314860	GAGGCGCGCAGGAAGGTGCTTGGGCTACTCCAATACCATCCCCAAACACCCCTGCTGATC	314919
Qy	64	ProLYsValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyG:nArgAsn	83
Db	314920	GAGACGCTCGAATGCCTGTGGTCGACGCT-----CAGAAACGCTCTG	314961
Qy	84	ValValValIThrGlu---GlyAsp:eulySyysValLeuAspGlyCysAlaproLeuthr	102
Db	314962	ATAAAGATCATCGAGACAGCGGCTTATCGCGCTCGTGGCGACAAC-----CAGACAGGG	315015
Qy	103	ArgPheThrAsnLYsLeuArgThrPheGLYArgThrPheThrGluLaTyRValaspPhe	122
Db	315016	CGGTTTCGAAGCCCGCTG-----	315033
Qy	123	CysIleAlaTyR:LysH:sLYsLeuProGlnLeuAsnAlaAlaGluLeuGlyIlePro	142
Db	315034	ATCCTGGCATCGACGGCGCGCTTCCTAGTTGGGCTCGAGCGGGCAGCTG---ATACCT	315090
Qy	143	AlaGluAspSerTyRLeuAla-----AlaAspPheLeu	153

```
Ds 315091 GCTCTTGGATCGCGTCTCGCGCGGTGACGCTCTTCTTCCTCCCTCTATCGCATCGCCTC 315150
Qy 154 GlyThrCysProLysLeuSerGlu-----LeuGlnGlnSerArgLys 167
Ds 315151 GAAGA*CTTCCACTCTAGCTGACACACTTCTCTACAAAGCGCTCGAAAG 315198

RESULT 13
US-09-252-991A-3735
; Sequence 3735, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3735
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3735

Alignment Scores:
Pred. No.: 0.694 Length: 1233
Score: 77.50 Matches: 60
Percent Similarity: 37.92% Conservative: 31
Best Local Similarity: 25.00% Mismatches: 90
Query Match: 7.82% Indels: 59
DB: 4 Gaps: 13

US-09-613-486-15 (1-198) x US-09-252-991A-3735 (1-1233)
Qy 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGly---ValAspLys 25
Ds 232 AGCGGATGACGGCTGGCGTCTCTCGCCAGCGAAGCTGAGCGGAAGTCCATTG 291
Qy 26 LysLeuLeuSerAlaGluValGluLysMetLeuValGlnLys----- 39
Ds 292 GTGATCTCAGCAGCGAGGTGGATCCATCTCGCCAGCGCCACCATTTCGATGTCGAG 351
Qy 40 -----GlyAlaProAsnGluGlyLeuValVal 49
Ds 352 TGCTTGGGCTCAATTTCTCGCGGACCTTGGCAAGCCA---TTCAGCCTGAGCGGATC 403
Qy 50 PheGlyLeuLeuLeuTyraLeuAlaAlaArgThrThrSerProLys---ValGlnArg 68
Ds 409 ACCGCCCTGTGACTCGCTACAACGGTCTGTCGCCAGGACCTCGCCGAGCAGATCGAGT 462
Qy 69 AlaAsp-----SerAspValIlePheSerAsnSerPheGlyGlu----- 81
Ds 469 GCGGAGTGCTCCCTCGCGGCGAGTGGTGGCGGCGCTCGCAACAGCGGATTCGAAGCC 528
Qy 82 -----ArgAsnValValValThrGluGlyAspLeu-----LysLysValLeuAsp 96
Ds 529 TACTACAGCCCAAGGTGGCCCTTGGATGGCGGCGGCTGATCGCGCCGAGGTCCTGGCA 588
Qy 97 -----GlyCysAlaProLeuThrArgPheThrAsnLysLeuArg 109
Ds 589 CGCTGGAAACACCGCATCTCGCGGTATTCGCCCGCTCGCATTTCTCTATGTATGGA 648
Qy 110 Thr-----PheGlyArgThrPheThrGluAlaTyraValAspPhe 122
Ds 649 ACCTACAACCTGTGTGCAAGAGCTGTTCTGGCAACTGTTTCAGCCAGGG----- 696
Qy 123 CysIleAlaTyraLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIlePro 142
Ds 697 ---CTGGCAGCAGGAGAGTGGCGCAGTGGGGCAGCCGATCAACCTGGCGTTCAAC 753
```

```
Qy 143 AlaGluAspSerTyraLeuAlaAlaAspPheLeu-----GlyThrCysProLysLeuSer 160
Ds 754 GTCCATCTCTCGCAACTGGGTTCGCCCTGGCGGAGAACATCTCGCGCTTGTCTGACC 813
Qy 161 Glu-----LeuGlnGlnSerArgLysMetPheAlaSerMetTyraAlaLeuLysThrGlu 178
Ds 814 GAGTTCACCTCGCGCCAGTTCGGTGAATGTCGAGATC-----ACCGAG 858
Qy 179 GlyGlyValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgGluValMet 198
Ds 859 ACCGGTGTGATCAGCGCTCTCTGCCACAGCAGCTGGAGAACCTGTGTCGTCTGGGAGCATG 918

RESULT 14
US-09-252-991A-3636/c
; Sequence 3636, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3636
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3636

Alignment Scores:
Pred. No.: 0.931 Length: 1491
Score: 77.50 Matches: 60
Percent Similarity: 37.92% Conservative: 31
Best Local Similarity: 25.00% Mismatches: 90
Query Match: 7.82% Indels: 59
DB: 4 Gaps: 13

US-09-613-486-15 (1-198) x US-09-252-991A-3636 (1-1491)
Qy 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGly---ValAspLys 25
Ds 101 AGCGGATGACGGCTGGCGTCTCTCGCCAGCGAAGCTGAGCGGAAGTCCATTG 1042
Qy 26 LysLeuLeuSerAlaGluValGluLysMetLeuValGlnLys----- 39
Ds 1041 GTGATCTCAGCAGCGAGGTGGATCCATCTCGCCAGCGCCACCATTTCGATGTCGAG 982
Qy 40 -----GlyAlaProAsnGluGlyLeuValVal 49
Ds 981 TGCTTGGGCTCAATTTCTCGCGGACCTTGGCAAGCCA---TTCAGCCTGAGCGGATC 925
Qy 50 PheGlyLeuLeuLeuTyraLeuAlaAlaArgThrThrSerProLys---ValGlnArg 68
Ds 924 ACCGCCCTGTGACTCGCTACAACGGTCTGTCGCCAGGACCTCGCGGAGCAGATCGAAGTC 865
Qy 69 AlaAsp-----SerAspValIlePheSerAsnSerPheGlyGlu----- 91
Ds 864 GCGGAGTGCTCCCTCGCGGCGAGTGGTGGCGGCTCGCAACAGCGGATTCGAAGCC 855
Qy 82 -----ArgAsnValValValThrGluGlyAspLeu-----LysLysValLeuAsp 96
Ds 804 TACTACAGCCCAAGGTGGCCCTTGGATGGCGGCGCTGATCGCGCCGAGGTCCTGGCA 745
Qy 97 -----GlyCysAlaProLeuThrArgPheThrAsnLysLeuArg 109
Ds 744 CGCTGGAAACACCGCATCTCGCGGTATTCGCCCGCTGCGCATTTCTCTATGTATGGA 685
```

```
QY 110 Thr-----PheGlyArgThrPheThrGluAlaTyrValAspPhe 122
|||
Db 684 ACCTACAACTGGTCGACAAAGCTGTTCTGGCAACTGTTTCAGCCAGGG- 637
QY 123 CysIleAlaTyrLysHisLysLysProGlnLeuAsnAlaAlaGluLeuGlyIlePro 142
|||
Db 636 ---CTGGCAGCCGCGAGGAGCTGGCGAGTGGGGCAGCCGATCAACTGGCGTTCAAC 580
QY 143 AlaGluAspSerTyrLeuAlaAlaAspPheLeu-----GlyThrCysProLysLeuSer 160
|||
Db 579 GTCCATCCCTCGCAACTGGGTTCGGCTGCCCTGGCCGAGAACATCTCGCGCTTGTGACC 520
QY 161 Glu-----LeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGlu 178
|||
Db 519 GAGTTCACCTCGCGCCCGAGTTCGGTGATGTTTCGAGATC-----ACCGAG 475
QY 179 G-GlyValValAsnThrProValSerAsnLeuArgGlnLeuGTYARGArgGluValMet 198
|||
Db 474 ACCGGGTTGATCAGCGCTCTCTSCACAGCAGCTGGAGAACCTCGTCGCTCGCGATCATG 415
```

## RESULT 15

```
US-09-252-991A-3824
; Sequence 3824, Application: US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubinfeld et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/274,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3824
; LENGTH: 4884
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3824
```

## Alignment Scores:

Pred. No.:	5.84	Length:	4884
Score:	77.50	Matches:	60
Percent Similarity:	37.92%	Conservative:	31
Best Local Similarity:	25.00%	Mismatches:	90
Query Match:	7.82%	Indels:	59
DB:	4	Gaps:	13

US-09-613-486-15 (1-198) x US-09-252-991A-3824 (1-4884)

```
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGly---ValAspLys 25
|||
Db 143 AGCGGATGACGGCGCTGGCGTTCCTTCGCCACGCAAGCTTAGCGGCAAGTCCATTGG 202
QY 26 LysLeuLeuSerAlaGluValGluLysMetLeuValGlnLys----- 39
|||
Db 203 GTGATCCTCAGCAGCGAGGTGGATCCATCTCGCCAGGCCACCATTTCGATGATCGAG 262
QY 40 -----GlyAlaProAsnGluGlyIleGluValVal 49
|||
Db 263 TGCGTGGCGCTCAATTTCTCGCGGACCTTGGCAAGCCA---TTCAGCTGGAGGGATC 319
QY 50 PheGlyLeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLys---ValGlnArg 68
|||
Db 320 ACCGCCCTGTTGATCTCGCTACAGCTCGTCGCCAGGACCTCGCGGACAGATCGAGTC 379
QY 69 AlaAsp-----SerAspValIlePheSerAsnSerPheGlyGlu----- 81
|||
Db 380 GCCGAGTGCCCTCGTGGGGAGCTGGTGGCGGCTCGACAAACGCGAGTTCGAAGCC 439
QY 82 -----ArgAsnValValValThrGluGlyAspLeu-----LysLysValLeuAsp 96
```

```
Db 440 TACTACCAAGCCAAAGTGGCCCTGGATGGCGCGCGCCTGATCGGCGCGAGGTCTCGCA 499
QY 97 -----GlyCysAlaProLeuThrArgPheThrAsnLysLeuArg 109
|||
Db 500 CGCTGGAAACCCCGCATCTCGCGTATTGCGCGCTCGCATTTCTCTATGTATGAA 559
QY 110 Thr-----PheGlyArgThrPheThrGluAlaTyrValAspPhe 122
|||
Db 560 ACCTACAACCTGGTCGACAAAGCTGTTCTGGCAACTGTTTCAGCCAGGG- 607
QY 123 CysIleAlaTyrLysHisLysLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIlePro 142
|||
Db 608 ---CTGGCAGCCGCGAGGAGCTGGCGAGTTCGGCGAGCCGATCAACTGGCGTTCAAC 664
QY 143 AlaGluAspSerTyrLeuAlaAlaAspPheLeu-----GlyThrCysProLysLeuSer 160
|||
Db 665 GTCCATCTCTCGCAACTGGGTTCGGCTGCCCTGGCGCGAGAACATCTCGCGCTGCTGACC 724
QY 161 Glu-----LeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGlu 178
|||
Db 725 GAGTTCACCTGGCGCCCGAGTTCGGTGA-TGTTTCGAGATC-----ACCGAG 769
QY 179 G-GlyValValAsnThrProValSerAsnLeuArgGlnLeuGlyArgGluValMet 198
|||
Db 770 ACCGGGTTGATCAGCGCTCTCTCGCAGCAGCTGGAGAACCTCGTCGCTCGCGATCATG 829
```

Search completed: November 8, 2003, 01:42:06

Job time : 140 secs

GenCore version 5.i.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 01:33:59 ; Search time 302 Seconds  
(without alignments)  
2092.085 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMSDSNLSNLVITDASSL.....GGWNTFVSNLRQLGRREV 195

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2141354 seqs, 1595479879 residues

Total number of hits satisfying chosen parameters: 4292708

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=c-frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USPFO.spool/US09613486/runat\_07112003\_120413\_27179/app\_query.fasta.i.391  
-DB=Published Applications NA -QFMT=fastap -SCFPI=nrnp -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -XATRI=X-blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUT=EMT=ptc -NOR=ext -HEAPSIZE=500 -WINLEN=3  
-MAXLEN=2000000000 -USER=US09613486 %CGN 1 1 221 @runat\_07112003\_120413\_27179  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOB.seq\*
- 2: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOB.seq\*
- 3: /cgn2\_6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/2/pubnpa/US06\_PUBCOB.seq\*
- 5: /cgn2\_6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOB.seq\*
- 7: /cgn2\_6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubnpa/US08\_PUBCOB.seq\*
- 9: /cgn2\_6/ptodata/2/pubnpa/US09\_PUBCOB.seq\*
- 10: /cgn2\_6/ptodata/2/pubnpa/US09\_PUBCOB.seq\*
- 11: /cgn2\_6/ptodata/2/pubnpa/US09C\_PUBCOB.seq\*
- 12: /cgn2\_6/ptodata/2/pubnpa/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/2/pubnpa/US10\_PUBCOB.seq\*
- 14: /cgn2\_6/ptodata/2/pubnpa/US10B\_PUBCOB.seq\*
- 15: /cgn2\_6/ptodata/2/pubnpa/US10\_NEW\_PUB.seq\*
- 16: /cgn2\_6/ptodata/2/pubnpa/US60\_NEW\_PUB.seq\*
- 17: /cgn2\_6/ptodata/2/pubnpa/US60\_PUBCOB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	83.5	8.4	640691	10 US-09-790-988-1	Sequence 1, Appli

2	79	8.0	1356	9	US-09-815-242-6624	Sequence 6624, Ap
3	79	8.0	1434	12	US-10-138-842A-15	Sequence 15, Appl
4	79	8.0	1434	12	US-10-039-112-15	Sequence 15, Appl
5	79	8.0	11597	10	US-09-070-927A-222	Sequence 222, App
6	79	9.0	536165	11	US-09-039-964-1	Sequence 1, Appli
7	77.5	7.8	665	10	US-09-074-300-5760	Sequence 5760, Ap
8	77	7.8	2091	12	US-10-032-585-6808	Sequence 6808, Ap
9	76.5	7.7	1068	14	US-10-156-761-4239	Sequence 4239, Ap
10	76.5	7.7	8095	10	US-09-080-107-2228	Sequence 2228, Ap
11	76.5	7.7	9025608	14	US-10-156-761-1	Sequence 1, Appli
c 12	75	7.7	600	9	US-09-864-761-8204	Sequence 8204, Ap
13	75.5	7.6	29729	10	US-09-070-927A-238	Sequence 238, App
14	75	7.6	81940	10	US-09-759-5083-1	Sequence 1, Appli
15	75	7.6	81940	12	US-09-073-119-724	Sequence 724, App
16	75	7.6	81940	12	US-09-060-706-1092	Sequence 1092, App
17	74.5	7.5	876	14	US-10-156-761-3299	Sequence 3299, Ap
18	74.5	7.5	1320	10	US-09-074-300-1589	Sequence 1589, Ap
c 19	74.5	7.5	3684	12	US-10-133-013-109	Sequence 109, App
c 20	74.5	7.5	3309400	10	US-09-738-626-1	Sequence 1, Appli
c 21	74	7.5	1161	10	US-09-954-531-575	Sequence 575, App
22	74	7.5	1612	14	US-10-205-823-376	Sequence 376, App
23	74	7.5	1715	9	US-09-925-301-543	Sequence 543, App
24	73.5	7.4	786	10	US-09-894-844-129	Sequence 129, App
25	73.5	7.4	1353	14	US-10-156-761-4968	Sequence 4968, Ap
c 26	73.5	7.4	2350	10	US-09-070-927A-109	Sequence 109, App
c 27	73	7.4	1020	10	US-09-038-842A-2260	Sequence 2260, Ap
28	73	7.4	1455	14	US-10-156-761-5889	Sequence 5889, Ap
29	73	7.4	5964	12	US-10-032-585-6528	Sequence 6518, Ap
c 30	73	7.4	49136	10	US-09-768-877-1	Sequence 1, Appli
c 31	73	7.4	59065	12	US-10-135-696-3	Sequence 3, Appli
c 32	73	7.4	9025608	14	US-10-156-761-1	Sequence 1, Appli
33	72.5	7.3	551	9	US-09-976-451-1	Sequence 1, Appli
c 34	72.5	7.3	3657	9	US-09-976-451-10	Sequence 10, Appl
35	72	7.3	1071	10	US-09-974-300-2722	Sequence 2722, Ap
36	72	7.3	1084	10	US-09-974-300-1197	Sequence 1197, Ap
37	72	7.3	1084	10	US-09-974-300-1308	Sequence 1308, Ap
38	72	7.3	2709	12	US-10-087-402-5	Sequence 5, Appli
39	72	7.3	3360	12	US-10-032-585-6333	Sequence 6333, Ap
c 40	72	7.3	5715	14	US-10-128-714-2237	Sequence 2237, Ap
c 41	72	7.3	5715	14	US-10-128-714-7237	Sequence 7237, Ap
c 42	72	7.3	5819	14	US-10-128-714-1237	Sequence 1237, Ap
c 43	72	7.3	5819	14	US-10-128-714-6237	Sequence 6237, Ap
c 44	72	7.3	7425	14	US-10-114-170-212	Sequence 212, App
c 45	72	7.3	7819	14	US-10-128-714-237	Sequence 237, App

ALIGNMENTS

RESULT 1  
US-09-790-988-1/c  
; Sequence 1, Application US/C9790988  
; Patent No. US20020127687A1  
; GENERAL INFORMATION:  
; APPLICANT: SHIGENOBU, SHUJI  
; APPLICANT: WATANABE, HIJEMI  
; APPLICANT: HATTORI, MASAHIRA  
; APPLICANT: SAKAKI, YOSHIOYUKI  
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APIIDS  
; FILE REFERENCE: 081356/0159  
; CURRENT APPLICATION NUMBER: US/09/790,988  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: JP2000-107160  
; PRIOR FILING DATE: 2000-04-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 640681  
; TYPE: DNA  
; ORGANISM: Buchnera sp.  
US-09-790-988-1

Alignment Scores: 3.46e+03 Length: 640681  
Pred. No.: 3.46e+03 Length: 640681

```
Score: 83.50 Matches: 32
Percent Similarity: 37.18% Conservative: 26
Best Local Similarity: 20.51% Mismatches: 53
Query Match: 8.43% Indels: 45
DB: 10 Gaps: 5

US-09-613-486-15 (1-198) x US-09-790-988-1 (1-640681)
QY 65 ysvValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsnVal 84
Db 435915 AAAAAATCAAGCTGGATCATTAAGCATTTGATGACGAAGGTACTCCAGCTCAAAAGAATATT 435856
QY 85 ValValThrGluGlyAspLeuLysLysValLeuAsp-----GlyCysAlaProLeuThrArg 103
Db 435855 TTAATTGAAATCGAATATATAAAATAATATGCAAGATAAACTTAATCGCGGTTTAATG 435795
QY 97 -----GlyCysAlaProLeuThrArg 103
Db 435795 GGTGTTAAGTCTACCGGAAATGTCGTCTGATCTTATTGCTGCTACCTATGCTCTGT 435736
QY 104 PheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyValAspPheCys 123
Db 435735 AGCATTAAT-----ACCATATG 435718
QY 124 IleAlaTyLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAla 143
Db 435717 TTATCTGGAAGTCTAAATAGATGATATAATTAAGAGGTGCACTATGGAATA----- 435664
QY 144 GAspSerTyLysLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGln 163
Db 435663 -----TATCTCTGAATTTTCTGGA-----GGACAGGTAGATATTACT 435625
QY 164 GlnSerArgLysValPheAlaSerMetTyAlaLeuLysThrGluGlyValValAsn 183
Db 435624 TCCGGAAATTTGTATTTCTACTTCGGAGCTTATTTAATTAATAAATAAGGAAATTTGT 435565
QY 184 ThrProValSerAsnLeuArgGlnLeuGlyArg-----ArgGluValMet 198
Db 435564 ACACCAATTAATAAACACTACTCTCAATAGGATCAGGATTAGAAGTTATG 435517

RESULT 2
US-09-815-242-6624
; Sequence 6624, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Christen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant C.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in:
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/197,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6624
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1356)
; US-09-815-242-6624

Alignment Scores:
Pred. No.: 0.851 Length: 1356
Score: 79.00 Matches: 53
Percent Similarity: 36.53% Conservative: 27
Best Local Similarity: 24.20% Mismatches: 73
Query Match: 7.97% Indels: 66
DB: 9 Gaps: 12

US-09-613-486-15 (1-198) x US-09-815-242-6624 (1-1356)
QY 1 MetGluLeuValSerAspSerAsnLeuSer-----AsnLeuValIleThrAspAla 17
Db 529 TTAGATTACGACAGAGGACTCGAATGCTGCGAGTGGATATGAACATCGTGATGACCGAGTCT 588
QY 18 SerSerLeu-----AsnGlyValAspLysLysLeuSerAlaGluVal 32
Db 589 GGTCAATTTGTGGAATTCAGGAAACGGGTGAGAGACTACCTTCTCAGGAGATGAGTTA 648
QY 33 GluLysMetLeuValGlnLysGlyAlaProAsnGluGlyLeGluValValPheGlyLeu 52
Db 649 AATCCCATGCTCTTTTAT-----GGAAAAACAGGATCGAGGAATTAATTCGCTAT 699
QY 53 -----LeuLeuTyAlaLeuAlaAlaArgThrThrSerProLysValGlnArgAla 69
Db 700 CAGAAAGACGCTTTATGCTTTAGCC-----CTGAGGAAGTCCCTCCCAA 747
QY 70 AspSerAsp-----ValIle-----Phe 75
Db 748 GATTCAAGAGGAAGTGTATGTCATTGCAACTAGAAATCTGTGTAAGCGAAAGACTTT 807
QY 76 SerAsnSerPheGlyGluArgAsrValValValThrGluGlyAspLeuLysValLeu 95
Db 808 AGTTCAATTTTGGTGAAAGAGGATATACAGTT-----AAGACGTTATTA 852
QY 96 AspGlyCysAlaProLeuThrArgPheThrAsn-----LysLeuArgThrPheGlyArg 113
Db 853 GAT-----TATCCGAATTTGCCAGATGTAGAGAAACTGGTCTGA 891
QY 114 ThrPheThrGluAlaTyValAspPheCysIleAlaTyLysHisLysLeuProGlnLeu 133
Db 892 ACATTGGAAGAA-----AATGCTGCTCTAAAGCTGAAACAAATT 930
QY 134 AsnAlaAlaAlaGluLeuGlyLeProAlaGluAspSerTyLeuAlaAlaAspPheLeu 153
Db 931 GCAGAAATTTTACAGAAACCCAGTATTAGCAGATGATTACAGGATTAATTTGTTGATGCGTTA 990
QY 154 GlyThrCysProLysLeuSerGluLeuGlnSerArgLysMetPheAlaSerMetTy 173
Db 991 GGGGAATGCGC-----GGAACTAT 1011
QY 174 AlaLeuLysThrGluGlyGlyValValValAsnThrProValSerAsnLeuArgGlnLeu 192
Db 1012 TCTGCTGTTTTGCTGCTGAGCCAACTAACGATGCTTCAAAATAATGCGAAACTACTA 1068

RESULT 3
US-10-138-842A-15
; Sequence 15, Application US/10138842A
; Publication No. US20030148390A1
; GENERAL INFORMATION:
; APPLICANT: GONSALVES, DENNIS
; APPLICANT: LING, KAI-SHU
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
; TITLE OF INVENTION: THEIR USES
```

FILE REFERENCE: 07678/025036  
CURRENT APPLICATION NUMBER: US/10/138-842A  
CURRENT FILING DATE: 2002-05-03  
PRIORITY APPLICATION NUMBER: US 09/579,259  
PRIORITY FILING DATE: 2000-05-25  
PRIORITY APPLICATION NUMBER: US 09/224,898  
PRIORITY FILING DATE: 1998-12-31  
PRIORITY APPLICATION NUMBER: US 08/770,544  
PRIORITY FILING DATE: 1996-12-20  
PRIORITY APPLICATION NUMBER: US 60/039,008  
PRIORITY FILING DATE: 1995-12-21  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: Fast-Seq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 1434  
TYPE: DNA  
ORGANISM: Grapevine Leafroll Virus  
US-10-138-842A-15

Alignment Scores:  
Pred. No.: 0.931 Length: 1434  
Score: 79.00 Matches: 43  
Percent Similarity: 40.49% Conservative: 40  
Best Local Similarity: 20.98% Mismatches: 60  
Query Match: 7.97% Indels: 62  
DB: 12 Gaps: 11

US-09-613-486-15 (1-198) x US-10-138-842A-15 (1-1434)

```
QY 12 LeuValIleThrAspAlaSerLeuAsnGlyValAspLysLysLeuSerAlaGlu 31
DB 904 TTCCGATACCGAAGCTCTACAGATAAAC-----GCAAGA 939
QY 32 ValGluLysMetLeuValGlnLysGly-----AlaProAsnGlu---GlyIleGlu 47
DB 940 CTACGCACACTCGTCTCTTCGAAGCGGAGTCAACACACGAGATATGGGGATATG 999
QY 48 ValValPheGlyLeuLeuLeuTyrAlaAlaAlaArgThrSerProLysValGln 67
DB 1000 ATAGTGGCCATGATACACTTTTCGTACTCTACTACTTAAGATATAACGGTCAAA 1059
QY 68 -----ArgAlaAspSerAspValIlePheSerAsnSerPheGlySerAsnVal 84
DB 1060 GACGGGTATAGGTGGAGACCGAATTA-----GCTCAAAAGAGAGTC 1101
QY 85 ValValThrGluGlyAspLeuLysLysValLeu-----AspGlyCysAlaPro 130
DB 1102 TACTTAAGTATTTCGAAGTAAGGAAGCTATATTAGGAGGAATACGGTGGCTCTCCA 1162
QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120
DB 1162 -----ACCAACACTGGCGATCTTCATGAGGTATTTTGTCTCACACCACTATC 1209
QY 121 AspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140
DB 1210 ACTCTACTATTAGAGAAGAAATTCAG---CCAGCGTGTACTGCCCTAGCTAAAGCAGCG 1266
QY 141 IlePro-----AlaGluAspSerTyrLeu 148
DB 1267 GTCCCGAAGAGGTTCACCTCGTACTGCTTCGACTCTCGACTACTGGATAACAGATATTAC 1325
QY 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnSerArgLysMet 168
DB 1327 CCGCGGAGCGGTG-----AAGGCTAACGCA 1353
QY 169 PheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrProValSerAsn 188
DB 1354 ATGGCTTCGGCTATAGCATTAATTAATCA-----GCTAAT 1386
QY 189 LeuArgGlnLeuGly 193
DB 1387 TTAAGCGTAAAGGT 1401
```

RESULT 4  
US-10-039-112-15  
Sequence 15, Application US/10039112  
Publication No. US2003C198942A1  
GENERAL INFORMATION:  
APPLICANT: Gonsalves, Dennis  
Ling, Kai-Shu  
TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS  
PROTEINS AND THEIR USES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 105;  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/039,112  
FILING DATE: 31-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60090008  
FILING DATE: 21-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/621  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 263-1304  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1434 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-10-039-112-15  
Alignment Scores:  
Pred. No.: 0.931 Length: 1434  
Score: 79.00 Matches: 43  
Percent Similarity: 40.49% Conservative: 40  
Best Local Similarity: 20.98% Mismatches: 60  
Query Match: 7.97% Indels: 62  
DB: 12 Gaps: 11  
US-09-613-486-15 (1-198) x US-10-039-112-15 (1-1434)  
QY 12 LeuValIleThrAspAlaSerLeuAsnGlyValAspLysLysLeuSerAlaGlu 31  
DB 904 TTCCGATACCGAAGCTCTACAGATAAAC-----GCAAGA 939  
QY 32 ValGluLysMetLeuValGlnLysGly-----AlaProAsnGlu---GlyIleGlu 47  
DB 940 CTACGCACACTCGTCTCTTCGAAGCGGAGTCAACACACGAGATATGGGGATATG 999  
QY 48 ValValPheGlyLeuLeuLeuTyrAlaAlaAlaArgThrSerProLysValGln 67  
DB 1000 ATAGTGGCCATGATACACTTTTCGTACTCTACTCTTAAGATATAACGGTCAAA 1059  
QY 68 -----ArgAlaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsnVal 84  
DB 1060 GACGGGTATAGGTGGAGACCGAATTA-----GCTCAAAAGAGAGTC 1101  
QY 85 ValValThrGluGlyAspLeuLysLysValLeu-----AspGlyCysAlaPro 100



1102 TACTTAAGTTATTCGGAAGTAAAGGAGCTATATTAGGAGGGAATACGCTCGCTCCA 1161  
101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120  
1162 -----ACCAACACTGTGCGATCCTTCATGAGGTATTTTCTCACACCACTATT 1209  
121 AspPheCysLeuAlaTyrCysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
1210 ACTCTACTTATACAGAGAAATTCAG---CCAGCGTGTACTGCCCTAGCTAAGCAGCGC 1266  
141 IlePro-----AlaGluAspSerTyrLeu 148  
1267 GTCCCGAAGAGGTTCACTCGCTTTCGACTTTCGACTACTGATAACAGATATTAC 1326  
149 AlaAlaAspPheLeuGlyThrCysPcLysLeuSerGluLeuGlnSerArgLysMet 168  
1327 CGGCGGAGCTGTGTG-----AAGCTAACCGCA 1353  
169 PheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrProValSerAsn 198  
1354 ATGGCTTTCGCTATACGATTATCA-----GCTAAT 1386  
189 LeuArgGlnLeuGly 193  
1387 TTAAGGCGTAAAGGT 1401

## RESULT 5

US-09-070-927A-222  
; Sequence 222, Application US/09070927A  
; Patent No. US20020120116A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Kunsch  
; Patrick J. Dillon  
; Steven Barash  
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Peptideptides  
; NUMBER OF SEQUENCES: 982  
; CORRESPONDENCE ADDRESS:  
; STREET: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070,927A  
; FILING DATE: 04-May-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/046,655  
; FILING DATE: 1997-05-16  
; APPLICATION NUMBER: 60/044,031  
; FILING DATE: 1997-05-06  
; APPLICATION NUMBER: 60/066,009  
; FILING DATE: 1997-11-14  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kenley K. Hoover  
; REGISTRATION NUMBER: 40,302  
; REFERENCE/DOCKET NUMBER: PB369  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 222:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11597 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 222:  
US-09-070-927A-222

Alignment Scores:  
Pred. No.: 26 Length: 11597  
Score: 79.00 Matches: 53  
Percent Similarity: 36.53% Conservative: 27  
Best Local Similarity: 24.20% Mismatches: 73  
Query Match: 7.97% Indels: 66  
DB: 10 Gaps: 12

US-09-613-486-15 (1-198) x US-09-070-927A-222 (1-11597)

QY 1 MetGluLeuMetSerAspSerAsnLeuSer-----AsnLeuValIleThrAspAla 17  
Db 2575 TTAGATTACCGAGAGGACTCGAATGCTGCGATGATGACATCGTGATGACCGAGTCT 2634  
QY 18 SerSerLeu-----AsnGlyValAspLysLysLeuLeuSerAlaGluVal 32  
Db 2635 GGTCGAATTTGTGGAATTCAGAGAACGGGTGAAGAACTACCTTCTCAGGAGATGAGTTA 2694  
QY 33 GluLysMetLeuValGlnLysGlyAlaProAsnGluGlyLeuValValPheGlyLeu 52  
Db 2695 AATGCCATGCTCTTTTA?-----GGAAGAACACGGGATTGAGGAATTAATTGCGTAT 2745  
QY 53 -----LeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLysValGlnArgAla 69  
Db 2746 CAGAACGAAGCGTTTATGTTTAGCC-----TCTGAGAGATCCCTCCCAA 2793  
QY 70 AspSerAsp-----ValIle-----phe 75  
Db 2794 GATTACAGAGAGAAAGTGATTGTCAATTGCGACAGAAATCTGCTAAAGCGAAAGAGTTT 2853  
QY 76 SerAsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLysValLeu 95  
Db 2854 AGTTCATTTTGTGTAAGAGGATACAGTT-----AAGACGTTTATTA 2898  
QY 96 AspGlyCysAlaProLeuThrArgPheThrAsn-----LysLeuArgThrPheGlyArg 113  
Db 2899 GAT-----TATCCGAATTTGCCAGATGTAGAGAACTGGTCTGA 2937  
QY 114 ThrPheThrGluAlaTyrValAspPheCysIleAlaTyrLysHisLysLeuProGlnLeu 133  
Db 2938 ACATTTCAAGAA-----AATGCTGCTCTAAAGAGCTGAAACAAATT 2976  
QY 134 AsnAlaAlaAlaGluLeuGlyIleProAlaGluAspSerTyrLeuAlaAlaAspPheLeu 153  
Db 2977 GCAGAAATTTTACAGAAACCCAGTACTAGCAGATGATTACAGGAATTAATTGTTGATGCCCTTA 3036  
QY 154 GlyThrCysProLysLeuSerGluLeuGlnSerArgLysMetPheAlaSerMetTyr 173  
Db 3037 GGAGGAATGCCA-----GGGAATTTAT 3057  
QY 174 AlaLeuLysThrGluGlyValValValAsnThrProValSerAsnLeuArgGlnLeu 192  
Db 3058 TCTGCTGTTTTGCTGTGAGCAACTAAGCATGCTTCAATAATATGCGAAACACTA 3114

## RESULT 6

US-09-939-964-1  
; Sequence 1, Application US/09939964  
; Publication No. US20030054522A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosenthal, Andre  
; APPLICANT: Freiberg, Christoph  
; APPLICANT: Perret, Xavier Philippe  
; APPLICANT: Broughton, William John  
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic  
; FILE REFERENCE: CARP0068  
; CURRENT APPLICATION NUMBER: US/09/939,964  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 09/214,808  
; PRIOR FILING DATE: 1999-06-22

```

RESULT: 7
US-09-974-300-576C
; Sequence 5760, Application US/39974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berkla, Randy M.
; APPLICANT: Clausen, Ib Grøth
; TITLE OF INVENTION: Methods For Monitoring
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: 60/279,526

```



TYPE: DNA  
ORGANISM: Streptomyces avermitilis



```
Db 13176 GCAGCGGTATGCGGGGKTTGAGTTACCTTTACTGTTCTCCAGCAACAGAGTCACT 13235
QY 34 LysMetLeuValGlnLysGlyAlaProAspGluGlyLeuValValPheGlyLeuLeu 53
Db 13236 CAGGAATAGTTGGTTATACCAAGAAATCTGACGTAGTCATTGGTGGGGCAGATT 13295
QY 54 LeuTyrAlaLeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAspVal 73
Db 13296 TTGGATGCTACGACTGCTGCTTTAGCTATTGTA-----GCAGGGCTGAAAT 13343
QY 74 IlePheSerAsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLysLys 93
Db 13344 ATTGTGATCCAGCTTCGATCGAAGAAACCGCTAAATAATTTCATTTATATCAAGTTCT 13403
QY 94 ValLeuAspGlyCysAlaProLeuThrArgPheThrAsnLysLeuArgThrPhe----- 111
Db 13404 TACTTGGCAGGCTGTATGACCATCACGAATGMAAGAACACTGMAAGCGGAGCAGAT 13463
QY 112 -----GlyArgThrPheThrGluAlaTyrValAspPheCysIleAla 125
Db 13464 ATTATTAAACTATTTCGGGAAGTGTATTGACCGAGTGTTATTTC-----GCA 13514
QY 126 TyrLysHsLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAsp 145
Db 13515 TTTAAAGCGCATTTGCCTCAGTTAAATATTATGCCAACGGGTGCTGTCAGTTTGGAAAT 13574
QY 146 -----SerTyrLeuAla 149
Db 13575 ATGCCCGAGTGTTCGTCGAGGTGTGACGGCGTCTGGTGTGGCAATCTGTAGCG 13634
QY 150 AlaAspPheLeuGlyThrCysProLysLeuSerGlnLeuGlnGlnSer---ArgLysMet 168
Db 13635 CCAGCAGCAACTGGGGATTTTGAAGAAAGTTCTGAGGTGCGCAAGCTATATGGAAATA 13694
QY 169 PheAlaSerMetTyrAlaLeuLysThrGluGlyValValAsnThrProValSerAsn 198
Db 13695 TTCCAACAATCAAGGGGTGTAGGCACATCGG-----GAA 13730
QY 189 LeuArgGlnLeuGlyArgArgGluVal 197
Db 13731 AGTCGTACATGGGAGAAATCATGTT 13757

RESULT 14
US-09-759-508B-1
; Sequence 1, Application US/09759508B
; Publication No. US20020182599A1
; GENERAL INFORMATION:
; APPLICANT: Fishman, Yark C.
; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease
; FILE REFERENCE: 00786/391C02
; CURRENT APPLICATION NUMBER: US/09/759,508B
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,787
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 81940
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (133)..(80910)
; OTHER INFORMATION:
US-09-759-508B-1
```

```
Alignment Scores:
Pred. No.: 2,22e+03 Length: 81940
Score: 75.00 Matches: 39
Percent Similarity: 37.71% Conservative: 27
Best Local Similarity: 22.29% Mismatches: 67
Query Match: 7.57% Indels: 42
DB: 10 Gaps: 8
```

```
US-09-613-486-15 (1-198) x US-09-759-508B-1 (1-81940):
QY 54 LeuTyrAlaLeuAlaAlaArgThrThrSer-----ProLys 65
Db 64723 STATACAGTGGAGCCAAATAATGCATCTGTTCTGCAAAAGCAGAAATAAAGTGAA 64782
QY 66 ValGlnArgAlaAspSerAspValIle-----PheSerAsnSerPheGlyGlu 81
Db 64793 GTACAGATACACAGCAAGAAAGTAGTTGGGCAATAGATTTCACCAATATTACTGGGAG 64842
QY 82 ArgAsnValValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeu 101
Db 64843 AAGATGACTCTGTGTGGATGCCCACTCAAT-----GACGGTTGGCTCCCAT 64893
QY 102 ThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAsp 121
Db 64894 ACCCACTACAATTTGAAGAAACCGGAAACCCAGCAGACTTGGCTGGGCACCTAATTGAGAT 64953
QY 122 PheCysIleAla-----TyrLysHis 128
Db 64954 AATGTGAAGCCCAAGTTACACTGCCATTAAACTAATAAACGGCAATGAATACCAATTC 65013
QY 129 LysLeuProGlnLeuAspAlaAlaGluLeuGlyIlePheAlaGluAspSerTyrLeu 148
Db 65014 CGTGTCTCTGAGTTTAAACAGTTTGGT---GTTGGCAGGCCACTTGATCTGATCCAGTG 65070
QY 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGlnLeuGlnGlnSerArgLysMet 168
Db 65071 GTTGTCTCAATACAAATATACCTGTTCTGATGCCCTGGCATTCAGAAACCTAGCAACATA 65120
QY 169 PheAlaSerMetTyrAlaLeu-----LysThrGluGlyGlyValValAsn 183
Db 65131 ACAGCAACAGCATTTACCTGCACATGGCAAGGCCAGCAATCAGATGGTGGC----- 65181
QY 184 ThrProValSerAsnLeuArgGln-----LeuGlyArgArgGlu 196
Db 65182 -----AGTGAATTCACACAGTATATCTCTTGAAGAGAGAGAA 65217

RESULT 15
US-09-873-319-724
; Sequence 724, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Kulkarni, William E.
; APPLICANT: Munger, Robert H.
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 724
; LENGTH: 81940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X90568
US-09-873-319-724
```

```
Alignment Scores:
Pred. No.: 2,22e+03 Length: 81940
Score: 75.00 Matches: 39
Percent Similarity: 37.71% Conservative: 27
Best Local Similarity: 22.29% Mismatches: 67
Query Match: 7.57% Indels: 42
DB: 12 Gaps: 8
```

Search completed: November 8, 2003, 03:17:34  
Job time : 2632 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: November 7, 2003, 14:47:03 ; Search time 21 Seconds  
(without alignments)  
906.733 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 MELNSDNLNLSNLTWTDASSL.....GGWNTFVSNLRQLGRREVM 195

Scoring table: BL0SUM62

Gapop 16.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	293	29.6	204	2	coat protein - sug
2	286	28.9	204	2	coat protein - sug
3	125	12.6	223	2	capsid protein p25
4	104	10.5	216	2	hypothetical prote
5	100.5	10.1	216	2	coat protein homol
6	100.5	10.1	240	2	capsid protein hom
7	99.5	10.0	216	2	hypothetical prote
8	92	9.3	1692	2	adenylate cyclase
9	90	9.1	419	2	P-450 monooxygenas
10	87	8.8	527	2	phosphoenolpyruvat
11	86.5	8.7	328	2	probable GDSJ-moti
12	86	8.7	360	2	probable GDSJ-moti
13	86	8.7	445	2	hypothetical prote
14	86	8.7	447	2	probable beta-lact
15	85.5	8.6	729	2	hypothetical prote
16	85	8.6	1013	2	type I restriction
17	84	8.5	784	2	transketolase faml
18	83.5	8.4	483	2	tidd protein (impo
19	83	8.4	530	2	ABC transporter, A
20	83	8.4	530	2	ABC transporter SSO
21	83	8.4	561	2	phosphoenolpyruvat
22	83	8.4	561	2	phosphoenolpyruvat
23	81.5	8.2	823	1	endopeptidase Ia (
24	81	8.2	299	2	purine nucleoside
25	81	8.2	317	2	ribokinase XF0366
26	81	8.2	714	2	probable 3-hydroxy
27	80.5	8.1	387	2	flagellin - Cherno
28	80.5	8.1	673	2	penicillin-binding
29	80	8.1	555	2	formate-tetrahydro

30	75.5	8.0	320	2	hypothetical prote
31	79	8.0	308	2	hypothetical prote
32	79	8.0	422	1	cytochrome P450 Cy
33	78.5	7.9	209	2	orotate phosphorib
34	77.5	7.8	392	2	probable two-compo
35	77.5	7.8	523	2	hypothetical prote
36	77	7.8	289	2	hypothetical prote
37	77	7.8	397	2	probable NADH oxid
38	77	7.8	538	2	probable sodium/hy
39	76.5	7.7	503	2	hypothetical prote
40	76.5	7.7	2712	2	hypothetical prote
41	76	7.7	283	2	(S)-2-hydroxy-acid
42	76	7.7	384	2	cyclin A-type (clo
43	76	7.7	835	2	vacuolar protein s
44	76	7.7	1016	2	translation initia
45	75.5	7.6	291	2	methyl-transferase

ALIGNMENTS

RESULT 1

S27900

coat protein - sugar beet yellows virus

C:Species: sugar beet yellows virus, SBV

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000

C:Accession: S27900; S34206

R:Brundstedt, J.; Moseley, J.; Hull, R.

submitted to the EMBL Data Library, August 1991

A:Description: Nucleotide sequence of cDNA encoding the coat protein of beet yellows vi

A:Reference number: S27898

A:Accession: S27900

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-204 <BRU>

A:Cross-references: EMBL:MS9452; NID:g323237; PID:AAA72955.1; PID:g323240

R:Agarovsky, A.A.; Koonin, E.Y.; Boyko, V.P.; Maiss, E.; Lunina, N.A.; Atabekov, J.G.

submitted to the EMBL Data Library, June 1993

A:Description: Complete nucleotide sequence of the Ukrainian isolate of Beet yellows vi

A:Reference number: S34202

A:Accession: S34206

A>Status: preliminary

A:Molecule type: Genomic RNA

A:Residues: 1-204 <AGR>

A:Cross-references: EMBL:X73475; NID:g313689; PIDN:CAA51861.1; PID:g313694

C:Superfamily: SBV probable coat protein

Query Match 29.6%; Score 293; DB 2; Length 204;

Best Local Similarity 37.6%; Pred. No. 6.6e-20;

Matches 71; Conservative 31; Mismatches 85; Indels 2; Gaps 2;

QY 7 SNLSNLVITDASSLNGVDKLLSNAEVEKMLVQKAPNEGIEVVFGLLYALAAARTTSKPV 66

DB 11 ATFNVSADOTCHGDCCKLRKNFECLKLGKVPDKLGLALGCLYNSCHITGSKNV 70

QY 67 GRADSDVIFNSF-GERNVVVTEGLKKVLDGGCAPLRF--NKLTFTGRTTEAYVDFCIA 125

DB 71 SVQPTSTFIKASFGGKELFLTHGELRSFLDSQKLLGKPKNKLRCFCRTFKQDYISPAKE 130

QY 126 YKHLKPLQNLAAELGIPAESYLAADFCTCPKLSLQSKRKFASMYALKTEGGVNTF 185

DB 131 YRGLPPIARNRHGLPAEDHYLAADFISTSTELTDIQOGRLLARENATHFTFS-SESP 189

QY 186 VSNLRQLGR 194

DB 190 VTSKQLGR 198

RESULT 2

S28715

coat protein - sugar beet yellows virus

C:Species: sugar beet yellows virus, SBV

C:Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 11-Jan-2000



```
C:\Accession: S28715
R:\Agronovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.
J. Gen. Virol. 72, 15-23, 1991
C:\Species: citrus tristeza clostovirus CTV
C:\Date: 11-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 11-Jan-2000
C:\Accession: D49804
R:\Pappu, H.R.; Karasev, A.V.; Anderson, E.J.; Pappu, S.S.; Hilf, M.E.; Febres, V.J.; Eck
S.M.; Dawson, W.O.; Lee, R.F.; Niblett, C.L.
Virology 199, 35-46, 1994
A:\Title: Nucleotide sequence and organization of eight 3' open reading frames of the cit
A:\Reference number: A49804; PMID:94160579; PMID:8116253
A:\Note: severe quick decline isolate 136
A:\Accession: D49804
A:\Status: preliminary
A:\Molecule type: genomic RNA
A:\Residues: 1-223 <PAP>
A:\Note: sequence extracted from NCBI backbone (NCBIN:144092, NCBI:P:144102)
C:\Superfamily: SBV probable coat protein

Query Match 12.6%; Score 125; DB 2; Length 223;
Best Local Similarity 25.9%; Pred. No. 0.00036;
Matches 56; Conservative 33; Mismatches 63; Indels 64; Gaps 10;

QY 18 SSLLYGVQKLLSA-EVERKMLVQKGA-----PNEGIEWFGLLLYALA 58
DB 29 SSVNLH:DPLITMDVRLST:QNAALNRD:FLTLKKGKHPNLPBKDKDHFIAVMYRLA 88
QY 59 ARTTS-----PKVQPAD---SDVIP-SNSFGERNVVVTEGDKKVLGDC 98
DB 89 VKSSSLQSDDDATGTYTREGVEVLSKLTWDVFNFSKIGNR----- 132
QY 99 APLTRFTNKLRTFGRTTTEAYVDVFCIAYKHKLQPLNAAALGIPAEEDSYLAADFLGTGCPK 158
DB 133 -----TNALRWGRINDALYLAF-CRNRNLSYGRPRLDAGIPAGHYLCADFUTGAG 184
QY 159 LSELQSRKQKFAFASMYALKTEGGVVNTVPSNLRQLGR 194
DB 185 LTDLECAVYIQAEQLLK-KRGADVVVNTVNRQLGK 219

RESULT 3
C49804
capsid protein p25 - citrus tristeza clostovirus CTV
C:\Species: citrus tristeza clostovirus CTV
C:\Date: 11-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 11-Jan-2000
C:\Accession: D49804
R:\Pappu, H.R.; Karasev, A.V.; Anderson, E.J.; Pappu, S.S.; Hilf, M.E.; Febres, V.J.; Eck
S.M.; Dawson, W.O.; Lee, R.F.; Niblett, C.L.
Virology 199, 35-46, 1994
A:\Title: Nucleotide sequence and organization of eight 3' open reading frames of the cit
A:\Reference number: A49804; PMID:94160579; PMID:8116253
A:\Note: severe quick decline isolate 136
A:\Accession: D49804
A:\Status: preliminary
A:\Molecule type: genomic RNA
A:\Residues: 1-223 <PAP>
A:\Note: sequence extracted from NCBI backbone (NCBIN:144092, NCBI:P:144102)
C:\Superfamily: SBV probable coat protein

Query Match 12.6%; Score 125; DB 2; Length 223;
Best Local Similarity 25.9%; Pred. No. 0.00036;
Matches 56; Conservative 33; Mismatches 63; Indels 64; Gaps 10;

QY 18 SSLLYGVQKLLSA-EVERKMLVQKGA-----PNEGIEWFGLLLYALA 58
DB 29 SSVNLH:DPLITMDVRLST:QNAALNRD:FLTLKKGKHPNLPBKDKDHFIAVMYRLA 88
QY 59 ARTTS-----PKVQPAD---SDVIP-SNSFGERNVVVTEGDKKVLGDC 98
DB 89 VKSSSLQSDDDATGTYTREGVEVLSKLTWDVFNFSKIGNR----- 132
QY 99 APLTRFTNKLRTFGRTTTEAYVDVFCIAYKHKLQPLNAAALGIPAEEDSYLAADFLGTGCPK 158
DB 133 -----TNALRWGRINDALYLAF-CRNRNLSYGRPRLDAGIPAGHYLCADFUTGAG 184
QY 159 LSELQSRKQKFAFASMYALKTEGGVVNTVPSNLRQLGR 194
DB 185 LTDLECAVYIQAEQLLK-KRGADVVVNTVNRQLGK 219

RESULT 4
C49804

C:\Accession: S28715
R:\Agronovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.
J. Gen. Virol. 72, 15-23, 1991
C:\Species: sugar beet yellows virus, SBV
C:\Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
C:\Accession: S27899
R:\Brunstedt, J.; Moseley, J.; Hull, R.
submitted to the EMBL Data Library, August 1991
A:\Description: Nucleotide sequence of cDNA encoding the coat protein of beet yellows vi
A:\Reference number: S27899
A:\Accession: S27899
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1-216 <BRU>
A:\Cross-references: EMBL:M59452; NID:g33237; PIDN:AAA72954.i; PID:g3323239
C:\Superfamily: SBV probable coat protein

Query Match 10.5%; Score 104; DB 2; Length 216;
Best Local Similarity 29.5%; Pred. No. 0.032;
Matches 41; Conservative 19; Mismatches 62; Indels 22; Gaps 6;

QY 61 TTPKQVORADSDVIFNSFGERNVVVTEGDKKVLGDCAPLTRFTNKLRTFGRTTTEAYV 120
DB 83 TTSTKVVYVGA---YEYTIQGGKFLVKDAWVFFLKECKMKFNKPNPVRTFCATFEDAY- 138
QY 121 DFCIAYKHKLQPLN---AAALGIPAEEDSYLAADFL---GTC-----PKLSELQSRKQKFA 170
DB 139 ---IVIARSLPKLNLNRCTIGKRGIPSGYFELGADFLTATSVCLNDHEKAIVLQASRAAID 195
QY 171 SMYALKTEGGVVNTVPSNLRQLGR 194
DB 196 RAVSSSVDDGKIV-----SLFDLGR 214

RESULT 5
C54235
coat protein homolog - sugar beet yellows virus
C:\Species: sugar beet yellows virus, SBV
C:\Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 10-Dec-1999
C:\Accession: S34205
R:\Agronovsky, A.A.; Koonin, E.V.; Boyko, V.P.; Maiss, E.; Lunina, N.A.; Atabekov, J.G.
submitted to the EMBL Data Library, June 1993
A:\Description: Complete nucleotide sequence of the Ukrainian isolate of Beet yellows vi
A:\Reference number: S34202
A:\Accession: S34205
A:\Status: preliminary
A:\Molecule type: genomic RNA
A:\Residues: 1-216 <AGR>
A:\Cross-references: EMBL:X73475; NID:g332699; PIDN:CAAS1860.i; PID:g332693
C:\Superfamily: SBV probable coat protein

Query Match 10.1%; Score 100.5; DB 2; Length 216;
Best Local Similarity 25.0%; Pred. No. 0.067;
Matches 50; Conservative 27; Mismatches 94; Indels 29; Gaps 7;

QY 5 SDSNLSNLVITDASSLNGVDKLLSAEVEKMLVQKGNIEGWFGLLLYALAARTTSP 64
DB 34 SEVNPKNLNRKEDELLGVIRERFKSELV-----ITDDEFVKHLAFALIRAAITTTST 86
QY 65 KYQRAADSDVIFNSFGERNVVVTEGDKKVLGDCAPLTRFTNKLRTFGRTTTEAYVDVFCI 124
DB 87 KVVNVVGA---YEYTIQGGKFLVKDAWVFFLKECKMKFNKPNPVRTFCATFEDAY----I 139
QY 125 AYKHKLQPL---NAAALGIPAEEDSYLAADFL---GTC-----PKLSELQSRKQKFA 174
DB 140 VIARSLPKLFLNRTIGKRGIPSGYFELGADFLTATSVCLNDHEKAIVLQASRAAIDRAVS 199
QY 175 LKTEGGVVNTVPSNLRQLGR 194
DB 200 SSVDDGKIV-----SLFDLGR 214

RESULT 6
C49804
```

```

capped protein homolog p27 - citrus tristeza closterovirus CTV
C:Species: Citrus tristeza closterovirus CTV
C>Date: 11-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 11-Jan-2000
C:Accession: C49804
S.M.; Dawson, H.R.; Karasev, A.V.; Anderson, E.J.; Pappu, S.S.; Hilf, M.E.; Febres, V.C.; Eck
R.; Pappu, H.R.; Karasev, A.V.; Lee, R.F.; Niblett, C.L.
Virology 199, 35-46, 1994
A:Title: Nucleotide sequence and organization of eight 3' open reading frames of the cit
A:Reference number: A49804; MUID:94160579; PMID:8116253
A:Accession: C49804
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-240 <PAP>
A:Cross-references: GB:U16304; GB:U02547; GB:120760; NID:9306738; PIDN:AACS9629.1; PID:9
A:Note: sequence extracted from NCBI backbone (NCBI:144092, NCBIP:144099)
A:Superfamily: SBV probable coat protein
Query Match 10.1%; Score 100.5; DB 2; Length 240;
Best Local Similarity 26.4%; Pred. No. 0.077;
Matches 57; Conservative 28; Mismatches 70; Indels 61; Gaps 13;
QY 2 ELMSSDNLNLCVVT-DASSNGVKKLLSAEVEKMLVQKAPNEGIEVVGGLYALAAAR 60
DB 61 EKTGSHLKVYVMTDFT-LENYKTKTEOLLVHLTIQK-----RLYTI--- 104
QY 61 TSPKQVQADSDVIFNSFGERNVVTGDLK-KVLDGCA---PLTRFT-----NKLRTE 111
DB 105 STSTKTKFRDQKCI-----SVQOGLRYKLLDKVVFFPIISKFDRETPTNALRKF 154
QY 112 GRFTTEAYVDFCIA-----YKHLQPLNAAAALG:PAEDSVLAADFL-GTCEPKLSE--- 161
DB 155 ACTFEELFJ---CMARLPDLYENK-----RTTRAGTPLKGLVSADFLSSLPVGYSEHER 207
QY 162 ----JQSRKMFASVYALKTEGGVNTPVSNLRQLR 194
DB 208 GILRASLSLA-----RRQGYEATELLNRLDQK 238

RESULT 7
S28714
Hypothetical protein 5 - sugar beet yellows virus
C:Species: sugar beet yellows virus, SBV
C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 10-Dec-1999
C:Accession: S28714
R.; Agronovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.
J. Gen. Virol. 72, 15-23, 1991
A:Title: Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA g
A:Reference number: S28710; MUID:9116305; PMID:1990061
A:Accession: S28714
A:Molecule type: DNA
A:Residues: 1-216 <AGR>
A:Cross-references: EMBL:X51462; NID:958878; PIDN:CAA17553.1; PID:958882
C:Superfamily: SBV probable coat protein
Query Match 10.0%; Score 99.5; DB 2; Length 216;
Best Local Similarity 25.0%; Pred. No. 0.083;
Matches 50; Conservative 27; Mismatches 94; Indels 29; Gaps 7;
QY 5 SDSNLSNLNVT-DASSLNGVKKLLSAEVEKMLVQKAPNEGIEVVGGLYALAAARTSP 64
DB 34 SVPNVKNLNKKTDELLGVIRFKSELV-----ITDEDFVKHLAPALIRANITTSV 86
QY 65 KVVQADSDV:FSNSFGERNVVTGDLK-KVLDGCAPLTRFTNKLRTFGRTTTEAYVDFCI 124
DB 87 KNYVGA---YETIGCKKFLVDKAWVFPLIKCKMKKFNPNVTFCATFEDAY----I 139
QY 125 AYKHLQPLQJ---NAAELGIP:PAEDSVLAADFL---GTC-----PKLSELQSRKMFASVYA 174
DB 140 VIARSLPKFLNRTIGKRGIPSGYEF-GADFLTATSVCLNDHKAIVLQASAAIDRAVS 199
QY 175 LKTEGGVNTPVSNLRQLR 194

```

```

DB 200 SSVDGKIV-----SLFDLGR 214

RESULT 8
A33988
adenylate cyclase (EC 4.6.1.1) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 31-Mar-2000
C:Accession: A33988; A33539; T39809
R.; Young, D.; Riggs, M.; Field, J.; Vojtek, A.; Broek, D.; Wigler, M.
Proc. Natl. Acad. Sci. U.S.A. 86, 7989-7993, 1989
A:Title: The adenylate cyclase gene from Schizosaccharomyces pombe.
A:Reference number: A33988; MUID:90046723; PMID:2682634
A:Accession: A33988
A:Molecule type: DNA
A:Residues: 1-1692 <YOO>
A:Cross-references: GB:M26699; NID:gl73338; PIDN:AAA35284.1; PID:gl73339
R.; Yamawaki-Kataoka, Y.; Tamacki, T.; Choe, H.R.; Tanaka, H.; Kataoka, T.
Proc. Natl. Acad. Sci. U.S.A. 86, 5693-5697, 1989
A:Title: Adenylate cyclases in yeast: a comparison of the genes from Schizosaccharomyce
A:Reference number: A33539; MUID:89345533; PMID:2668944
A:Accession: A33539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1692 <YAM>
A:Cross-references: GB:M24942; NID:gl73378; PIDN:AAA35301.1; PID:gl73379
R.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gailardin, C.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z21981
A:Accession: T39809
A:Status: preliminary; translated from GB/EMBL/DDB
A:Molecule type: DNA
A:Residues: 1-1692 <LYN>
A:Cross-references: EMBL:AL023859; PIDN:CAA19571.1; GSPDB:GN00367; SPDS:SPRCL9C7.03
A:Experimental source: strain 972h-; cosmid c19C7
C:Genetics:
A:Map position: 2
A:Gene: SPBC19C7.03
A:Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homolo
C:Keywords: phosphorus-oxygen lyase
F:1328-1413/Domain: yeast adenylate cyclase catalytic domain homology <YACC>
Query Match 9.3%; Score 92; DB 2; Length 1692;
Best Local Similarity 24.7%; Pred. No. 5.9;
Matches 42; Conservative 26; Mismatches 68; Indels 34; Gaps 7;
QY 27 LLSAEVKMLVQKAPNEG:EVVFGLLLYALAAARTTSPKQVQADSDV:FSNSFGERNVV 86
DB 339 LQFENTERILL---PHEQPCIIIFERLLSLFGCKVTS-----DEEINEDNYSVARLVF 388
QY 87 TEGDLKKVLDGCAPLTRFTNKLRT---FQRTTEAYVDFCIAYKHLQPLNAAAALG:IP 142
DB 389 TTMDI-----GADVLRKFSEKKITANLDIRSNLEVIPVK:YPAHAEL:SLNVSHNLSLD 443
QY 143 AEDSYLAADF:GTCPLKSELQSRKMFASVYALKTEGGVNT:TPVSNLRQL 192
DB 444 ----LFLDFMERCVKLRJDISNN-----LRSFRG---KPIALRQL 478

RESULT 9
A33246
P-450 monooxygenase virH (imported) - Agrobacterium tumefaciens (strain C58, Dupont) P
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: A33246
R.; Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Katayama, T.; Levy, R.; Li, M.; Mc Clel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

```

A:Reference number: AB2577; MUID:21605550; PMID:11743193

A:Accession: AB2577

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-419 <K>

A:Cross-references: GB:AE008690; PIDN:AAL46386.1; P.D:G:7744179; GSPDB:GN00169

A:Experimental source: strain C58 (Dipert)

C:Genetics:

A:Gene: virH

A:Genome: plasmid

C:Superfamily: Agrobacterium plasmid cytochrome P450 pinF1; cytochrome P450 homology

C:Keywords: heme; iron; metalloprotein

P:366/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 9.1%; Score 90; DB 2; Length 419;

Best Local Similarity 25.1%; Pred. No. 1.5;

Matches 52; Conservative 26; Mismatches 71; Indels 58; Gaps 11;

QY 20 LNVGDKLLSAEVEKMLVQKAPNEGIE--VVFGLLYALAAARTTSPKQVQADSDVFSN 77

DB 67 LIGTPRTQITETMLM-----NGVTRGAVDFDIRYMLFSNGEVHVKRSA---PAK 117

QY 78 SFGERNVVVTEGLKXVLDGCA--LTRFTNKLRTFGRTFTAYVDFCIAYKHKLPQLNAA 136

DB 119 TEAFR-----MIDALRPEITKLTLEHLDVQVAVDD--PDFAEWYASKLPALTA 164

QY 137 RELGIPAE-----SYLAADFLGTCPLSELO-----OSRKM-- 168

DB 165 SVLGPFPGDAPFTFLVNVSCSPSGWGEDFP--EASAVELOYDVRVAVDRSRISD 224

QY 169 -PASMY--ALKTEGGVWNTPVSNLRLQ 192

DB 225 DFLCYLKAVREGTL--SPIEIVQL 249

RESULT 10

F69673

C:phosphoenolpyruvate carboxykinase pckA - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000

C:Accession: F69673

R:Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A:Authors: Kunat, F.; Cgasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

A, Map position: 2  
C, Superfamily: myrosinase-associated protein MyAP

Query Match 8.7%; Score 86; DB 2; Length 360;  
Best Local Similarity 26.5%; Pred. No. 2.9;  
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLVITDASSLNGV--DKLLSAEVEKMLVQKAPNEG1---EVVFGLLLYALAAAR 60  
DB 137 SQOPNMFKSVIARLKGIVGCKKAXEILINNAFVVSNGPNCDFILNYEIPSRLEVPFISG 196  
QY 61 TTSPPKVRADSDVIFSNFSGERNVWVTEGLKKVLDGCAPLTRFNKLRFGKTFTEAYV 120  
DB 197 YQDFILKRLNFVRESYSLGVANLV--GGLPM--GCUPI--HMTAKPFIREFCLEHNN 251  
QY 121 DRCIAYKHKPOLNAAE--GIPAEYSYLAADFLGCPKSELQOSRKMPASMYALKTEG 179  
DB 252 KQSVLYNEKLQNLJPOIEASJFG-SKFLYADVYN--PMKEMIQNFSK-----YGFKEIKR 303  
QY 180 GVVNT 184  
DB 304 GCCGT 308

Query Match 8.7%; Score 86; DB 2; Length 447;  
Best Local Similarity 26.0%; Pred. No. 3.9;  
Matches 40; Conservative 15; Mismatches 41; Indels 58; Gaps 8;

QY 26 KLLSAEVEKMLVQKAPNEGIEVVFGLLLYALAAARTTSPKVRADSDVIFSNFSGERNV- 84  
DB 5 KLESFIVEKMAERK-VFG-SISII-----KQGDVVYAKGFGYERVE 44  
QY 35 -----VTEGLKKVLDGCAPLTRFN- KLRFGKTFTE 116  
DB 45 ARLPSTPETTYGIGSITKSTFTALAIKMLVEGGLS--LDD--PVEKFN--KLRFGGSPVT 100  
QY 117 EAYVDFCIAYKHKPOLNAAEEL--GIPAEOSYL 148  
DB 101 ---VHLLTHSSGIFSLGYAEAFIDGMVGGONWL 131

RESULT 15  
T08639  
hypothetical protein b2627 - Escherichia coli (strain K-12)  
C, Species: Escherichia coli  
C, Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 01-Mar-2002  
C, Accession: T08639; E65041  
R, Plunkett, G.  
submitted to the EMBL Data Library, September 1995  
A, Reference number: Z16465  
A, Accession: T08639  
A, Status: preliminary; translated from GB/EMBL/DBJ  
A, Molecule type: DNA  
A, Residues: 1-729 <PLU>  
A, Cross-references: EMBL:U36840; NID:gl033110; PID:g1033122  
A, Experimental source: strain K12, substrain MGL655  
R, Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A, Title: The complete genome sequence of Escherichia coli K-12.  
A, Reference number: A64720; MUID:97426617; PMID:9278503  
A, Accession: E65041  
A, Status: preliminary; nucleic acid sequence not shown; translation not shown  
A, Molecule type: DNA  
A, Residues: 1-729 <BLAT>  
A, Cross-references: GB:AE000348; GB:U00096; NID:gl788975; PIDN:AAC75675.1; PID:g788980  
A, Experimental source: strain K-12, substrain MGL655  
C, Genetics:  
A, Map position: 57 min

Query Match 8.6%; Score 85.5; DB 2; Length 729;  
Best Local Similarity 22.4%; Pred. No. 8.1;  
Matches 41; Conservative 31; Mismatches 64; Indels 47; Gaps 8;

QY 2 ELMDSNLSNLVI---TDASSLNGVDKLLSAEVEKMLVQKAPNEGIEVVFGLLLYALAA 59  
DB 520 EMDNDEFFSLVFHMTATNPLTKFAKIIARLPHTSRNGVPVKPTDVMIAKLAGYLSA 579  
QY 60 RTTSPKVRADSDVIFSNFSGERNV--VTEGLKKVLDGCAPLTRFNKLRFGKTF 114  
DB 580 ESYSEYLK---NQIDYARQWISGEKRTLSIALNNDUKL-----TN---TFGYT 623  
QY 115 FTEA-----YVDFCIAYK--HKLPQNLAAAEELGIPAEOSYLAA 150

A, Map position: 2  
C, Superfamily: myrosinase-associated protein MyAP

Query Match 8.7%; Score 86; DB 2; Length 360;  
Best Local Similarity 26.5%; Pred. No. 2.9;  
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLVITDASSLNGV--DKLLSAEVEKMLVQKAPNEG1---EVVFGLLLYALAAAR 60  
DB 137 SQOPNMFKSVIARLKGIVGCKKAXEILINNAFVVSNGPNCDFILNYEIPSRLEVPFISG 196  
QY 61 TTSPPKVRADSDVIFSNFSGERNVWVTEGLKKVLDGCAPLTRFNKLRFGKTFTEAYV 120  
DB 197 YQDFILKRLNFVRESYSLGVANLV--GGLPM--GCUPI--HMTAKPFIREFCLEHNN 251  
QY 121 DRCIAYKHKPOLNAAE--GIPAEYSYLAADFLGCPKSELQOSRKMPASMYALKTEG 179  
DB 252 KQSVLYNEKLQNLJPOIEASJFG-SKFLYADVYN--PMKEMIQNFSK-----YGFKEIKR 303  
QY 180 GVVNT 184  
DB 304 GCCGT 308

Query Match 8.7%; Score 86; DB 2; Length 445;  
Best Local Similarity 24.7%; Pred. No. 3.8;  
Matches 54; Conservative 41; Mismatches 80; Indels 44; Gaps 12;

QY 1 MELMSNLSNLVITDAS-----SLNGVDKLLSAEVEKMLVQKAPNEGIEV----- 48  
DB 32 IKILNDSNLSVLIVDTNGKFSVIGNLGAENKLGEGVEMENLKIIDNSGGIAIPGFVD 91  
QY 49 -----VF-GLLLYALA---ARTSPKVRADSDVIFSNFSGERNVWVTEGLKKVLDGCA 99  
DB 92 GHSHPVFGSDRHFPAKLAGATYMEVQAGGIFTTN---KTRASSQDKRDFEELA 148  
QY 100 PLTRFNKLRFGKTFTEAYVDFCIAYKHKPOLN--AAAEELGIPAEOSYLAACFLG--T 155  
DB 149 -----KMLRSGLTTLEAKSGVGLNVDEXKMLRVLATENPNIPJE-----VSATCGAHA 199  
QY 156 CPKLS-ELOQSR-----KMFASMYALKTEGVVNTPVSNL 189  
DB 200 VPKGSYEYQTRMICEELIPKTEDEKRNGLKN--VENI 236

RESULT 14  
G75201  
probable beta-lactamase (EC 3.5.2.6) PAB0087 - Pyrococcus abyssi (strain Ocsay)  
N, Alternate names: penicillin-binding protein homolog  
C, Species: Pyrococcus abyssi  
C, Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 05-Nov-1999  
C, Accession: G75201

624 LKVL\$MEDVVKH\$H\$V\$K\$G\$R\$S\$K\$V\$D\$Y\$T\$H\$V\$K\$L\$A\$F\$E\$F\$H\$U\$P\$G\$W\$A\$L\$E\$B\$G\$P\$Q\$T\$L\$R\$J\$V 683  
QY 15: DFL 153  
Db 684 DLL 686

Search completed: November 7, 2003, 14:50:22  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: November 7, 2003, 14:45:43 : Search time 17 seconds  
(without alignments):  
547.723 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991

Sequence: 1 MELMSDSNLSNDVTDASSL.....GGVNTVPVN...PQLGREVV. 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127963 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 3

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	12.6	223	1 COAT_CTV36	Q06866 citrus tris
2	92	9.3	1692	1 CYAA_SCHPO	P14605 schizosacch
3	89.5	9.0	397	1 APX_YARLI	Q92389 yarrowia li
4	87	8.8	527	1 PPKC_BACSU	P54418 bacillus su
5	85.5	8.6	729	1 YFK_ECOL	P52126 escherichia
6	84	8.5	537	1 PPKC_RHOP	Q92nh4 rhodospirillum
7	83.5	8.4	483	1 TLDD_BUCAI	P57478 buchera ap
8	83	8.4	395	1 ARGU_LEPIK	Q8eyv8 l arginine
9	83	8.4	336	1 PPKC_AGRTS	Q8uj94 agrobacteri
10	81.5	8.2	431	1 ACKA_THETN	Q8r9v4 thermoanaer
11	81.5	8.2	813	1 LON2_BORBU	O51558 borelia bu
12	80	8.1	366	1 SP5V_MOUSE	P97355 mus musculu
13	79	8.0	308	1 OTCA_SINY3	Q55497 synectocyst
14	79	8.0	422	1 CPKC_AGRU	P24466 agrobacteri
15	79	8.0	527	1 PPKC_FUSNN	Q8rei2 fusobacteri
16	79	8.0	536	1 PPKC_RHIME	P43085 rhizobium m
17	79	8.0	609	1 Y4PA_RHISN	P55610 rhizobium s
18	78.5	7.9	209	1 PYRE_LACLA	Q9cgm8 lactococcus
19	78	7.9	477	1 DL22_PSEPU	P31052 psuedomonas
20	76.5	7.7	500	1 HI22_RHILC	Q98ep7 rhizobium l
21	76.5	7.7	641	1 CAN6_RAT	O88501 rattus norv
22	76	7.7	1016	1 IF2C_ARATH	Q9shil arabidopsis
23	75.5	7.6	538	1 PPKC_BRAJA	Q89bk7 bradyrhizob
24	75	7.6	277	1 ARGE_LFPBI	P13440 leptospira
25	75	7.6	536	1 PPKC_BRUME	Q8ye41 bruceella me
26	75	7.6	536	1 PPKC_BRUSE	Q8fy05 bruceella me
27	75	7.6	683	1 CLPB_SINP7	P53533 synectococc
28	74.5	7.5	222	1 Y175_METJA	Q57639 methanococc
29	74.5	7.5	359	1 HIS8_SALT	Q82519 salmoneilla
30	74.5	7.5	434	1 PURA_WIGBR	Q84322 wigleswort
31	74.5	7.5	581	1 AMY1_SCHPC	Q09840 schizosacch
32	74	7.5	366	1 SP5V_HUNAN	P52788 homo sapien
33	73.5	7.4	262	1 YY00_MYCTU	Q50725 mycobacteri

```

34 73.5 7.4 407 1 PEPT_STRPY
35 73.5 7.4 657 1 TKT_RHOSH
36 73.5 7.4 707 1 ORC1_SCHPO
37 73.5 7.4 862 1 SLA2_BACAN
38 73.5 7.4 1051 1 AP5B_EMENI
39 73 7.4 198 1 KAD_RHIL0
40 73 7.4 486 1 YC24_ODOSI
41 73 7.4 532 1 SAT1_MOUSE
42 73 7.4 886 1 RAS0_ARCFU
43 73 7.4 1024 1 U203_ARATH
44 72.5 7.3 278 1 PROC_VIBAL
45 72.5 7.3 496 1 AMYB_TR-RP

```

## ALIGNMENTS

```

RESULT 1
COAT_CTV36
ID COAT_CTV36 STANDARD; PRT; 223 AA.
AC CC5686;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Coat protein.
OS Citrus tristeza virus (isolate T36) (CTV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC Closterovirus.
OX NCB1_TaxID=31712;
RN 1;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-38.
RX MEDLINE=91237334; PubMed=2033386;
RA Sekiya M.E., Lawrence S.D., McCaffery M., Cline K.;
RT "Molecular cloning and nucleotide sequencing of the coat protein gene
of citrus tristeza virus.";
RL J. Gen. Virol. 72:1013-1020(1991).
RN 2;
RP SEQUENCE FROM N.A.
RX MEDLINE=94160579; PubMed=8116253;
RA Pappu H.R., Karasev A.V., Anderson E.J., Pappu S.S., Hilf M.E.,
RA Febres V., Eckloff R.M.G., McCaffery M., Boyko V., Gowda S.,
RA Dolja V.V., Koonin E.V.;
RT "Nucleotide sequence and organization of eight 3' open reading frames
of the citrus tristeza closterovirus genome.";
RL Virology 199:35-46(1994)
CC -!- PRT: CONSISTS OF AT LEAST TWO SIZE VARIANTS, CP1 AND CP2, WHICH
CC RESULT OF POST-TRANSLATIONAL PROTEOLYSIS AT SITES APPROXIMATIVELY
CC 12 TO 15 AND 26 AA FROM THE N-TERMINUS RESPECTIVELY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch.
-----
EMBL; M76485; AAA42924.1;
EMBL; U16304; AAC59630.1;
InterPro; IPR002679; Closter_coat.
Pfam; PF01785; Closter_coat; 1.
KW Coat protein.
SQ SEQUENCE 223 AA; 24909 MW; 00493F3D330220BB CRC64;

```

```

Query March 12.6%; Score 125; DB 1; Length 223;
Best Local Similarity 25.9%; Pred. No. 0.00025;
Matches 56; Conservative 33; Mismatches 63; Indels 64; Gaps 10;

QY 18 SSLN-GVDEKLLSA-EVEKMLVQKGA-----PNEGIVVFGLLLYALA 58
||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 29 SSVNLHIDPLTILMNDVRLSTQNALNRLFLTLKGKHPNLFDKOKFRIAMLYRLA 58
||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 59 ARTTS-----PKVQRAD---SDVIF-SNSFGERNVVVTGDLKKVLDCG 98

```

```

132 89 VKSSLSQSDDDATGITYTREGVEVLSOKLWTDVFNKSGIGNR----- 132
133 99 APLTRFTNKLRTFGRTTETAYVDFCIAYKHKLPLQNLAAAEIGIPAECDYSYLAADFGLGTCPK 158
134 133 -----TNAIRVGRGTNDALYLAFQ-RQNRNLISYGRPLDAGIPAGYHYLCDFL-TGAG 184
135 159 LSELOQRKXKFAFMALXTEGGVWVTFVSNLRQLQR 194
136 185 LTDLECAVYIQAEOLK-KRGADCVVTVNVAQLGK 219
RESULT 2
CYAA SCHPO STANDARD; PRT; 1692 AA.
AC PL4605;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl
DE cyclase).
CYRI OR SPC19C7.03.
GN Schizosaccharomyces pombe (fission yeast).
OS Schizosaccharomycetes;
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1];
RP SEQUENCE FROM N.A.;
MEDLINE=90046723; PubMed=2682634;
Yamawaki-Kataoka Y., Field J., Vojtek A., Broek D., Wigler M.;
"The adenylyl cyclase gene from Schizosaccharomyces pombe.";
Proc. Natl. Acad. Sci. U.S.A. 86:7989-7993(1989);
[2];
RP SEQUENCE FROM N.A.;
MEDLINE=89345333; PubMed=2668944;
Yamawaki-Kataoka Y., Tanacki T., Choe H.-R., Tanaka H., Kataoka T.;
"Adenylyl cyclases in yeast: a comparison of the genes from
Schizosaccharomyces pombe and Saccharomyces cerevisiae.";
Proc. Natl. Acad. Sci. U.S.A. 96:5693-5697(1999);
[3];
RP SEQUENCE FROM N.A.;
STRAIN=972;
MEDLINE=21848401; PubMed=11859360;
Wood V., Williams R., Rajandream M.A., Lyne V., Lyne R., Stewart A.,
Sgouros J., Peat N., Hayles C., Baker S., Basham D., Bowman S.,
Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
Gentles S., Goble A., Hamlin N., Harris D., Hidergott J., Hodgson G.,
Hollroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagers K.,
James K., Jones D., Jones M., Leather S., McDermid S., McLean J.,
Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
O'Neil K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Woodward J., Volkhardt G., Aert R., Robben C., Grimprez B.,
Wetjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
Gabel C., Fuchs M., Fritz C., Holzer E., Voestl D., Hilbert H.,
Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
Domínguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.E.,
Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
"The genome sequence of Schizosaccharomyces pombe.";
Nature 415:871-880(2002);
-!- FUNCTION: Plays essential roles in regulation of cellular
metabolism by catalyzing the synthesis of a second messenger,
cAMP.
```

```

87 TEGDLKKVLDGCAPIRTRETNKLT-----PGRFTTEAYVDFCIAYGKLPQANAAELCIP :42
389 TMDI-----GADVLRKSEKITTANLDIRSNLENLVPKVIYPYAHLEISUNVSHLSLD 443
143 AEDSYLAADFGLCTCPKSELQSRKXFASMYALKTEGGVNTPTVSNRLQL 192
444 -----LPLDMERCVKLKELDISNN-----LRSPRG---KPIFAJRL 478

RESULT 3
ID_XP_YARLI STANDARD; PRT: 397 AA.
AC Q92389;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acid extracellular protease precursor (EC 3.4.23.-).
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=148;
RC MEDLINE=97039861; PubMed=8885467;
RA Young T.W., Wadeson A., Glover D.J., Quincey R.V., Barlin M.J.,
RA Kamei E.A.;
RA "The extracellular acid protease gene of Yarrowia lipolytica:
RA sequence and pH-regulated transcription.";
RA Microbiology 142:2913-2921(1996);
RA -- SUBCELLULAR LOCATION: Secreted.
RA -- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
RA
RA This SWISS-PROT entry is copyrighted. It is produced through a collaboration:
RA between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See http://www.isb-sib.ch/announce/
RA or send an email to license@isb-sib.ch).
RA
RA EMBL; X97068; CAAG5778.1; --
RA HSP: P07339; 1LYB.
RA MEROPS; A01.036; --
RA InterPro; IPR001969; Asparticase site.
RA InterPro; IPR001461; AsparticaseA1.
RA Pfam; PF00026; asp; 1.
RA PRINTS; PR00792; PPSIN.
RA PROSITE; PS00141; ASP_PROTEASE; 1.
RA Hydrolase; Aspartyl protease; Zymogen; Signal.
RA SIGNAL 1 17
RA PROPEP 19 ?
RA CHAIN ? 397
RA ACT_SITE 77 77
RA ACT_SITE 264 264
RA DISULFID 93 100
RA DISULFID 302 343
RA CARBOHYD 88 88
RA CARBOHYD 310 310
RA CARBOHYD 314 314
RA SEQUENCE 397 AA; 4208. MW; BD678814B4B8984F CRC64;
RA
RA Query Match 9.0%; Score 89.5; DB 1; Length 397;
RA Best Local Similarity 21.9%; Pred. No. 0.91;
RA Matches 46; Conservative 37; Mismatches 86; Indels 41; Gaps 8;
RA
RA 15 TDASSLGVYKXLLSAEVEKMLVQKQ-APNEG:EVVFGLLVALAARTSPKVGKAD--S 71
RA 136 TNAVIGGA-----SAPGKFGVNGDLSGGFSVMVFGVGNDSADTS:SAQLQKSGEIS :90
RA 72 DVIFNSFGERNVVVTEGSLKKVLGDCAPLRTFTNKLRFTFGKTFTE-AYVDFFC:AYKHKL 130

```

```

191 RNLYGMSFSDANLAGTNDNSEITFGAINTGRTGSLKTIPTVATQGGYOHFVSASGKF 250
131 PQJ-----NAAAEIGIPAEYSYLAAD-----PLG-----TCPKJSELQ 163
251 GDVDLFDNDLVLDGSGTTTYLKSDDYNAFLGGLEDLIDTLSDYSGWHGPGCSENSKIN 310
164 -----OSRKMFAFMYALKTEGGVNTPTVSN 188
311 PTYNFSGKEITVTGHDLAIPGNVAVNSVDOS 340

RESULT 4
ID_PCK_BACSU STANDARD; PRT: 527 AA.
AC P54418; C34304;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49); (PEP
DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
OX PKA OR PPC.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=423;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RA "Sequencing and functional annotation of the Bacillus subtilis genes
RA in the 203 kb rnb-dnaB region.";
RA Microbiology 143:3431-3441(1997);
RA [2]
RA SEQUENCE FROM N.A.
RA STRAIN=168;
RA MEDLINE=9804033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Besieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codari J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Huilo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nockack M.,
RA Noone L., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
RA subtilis.";
RA Nature 390:249-256(1997);
RA [3]
RA SEQUENCE OF 1-165 FROM N.A.
RA STRAIN=168 / PY79;
RA MEDLINE=96345628; PubMed=875891;
RA Yocum R., Perkins J.B., Howitt C.L., Pero J.;

```



"Cloning and characterization of the metE gene encoding S-adenosylmethionine synthetase from *Bacillus subtilis*.";  
J. Bacteriol. 178:4604-4610(1996).  
-!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate + CO(2).  
-!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
-!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP) family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: AF008220; AAC00377.1; -  
EMBL: Z99119; CAB15034.1; -  
EMBL: U52812; AAB17065.1; -  
PIR: F69673; F69673.  
HSSP: P22259; 1AYL.  
Subtilist: BG11841; pckA.  
HAMAP: MF\_00453; -; 1.  
InterPro: IPR001272; PEPC\_K\_ATP.  
Pfam: PF01293; PEPC\_K\_ATP; 1.  
ProDom: PD004723; PEPC\_K\_ATP; 1.  
TIGRFAMs: TIGR00224; pckA; 1.  
PROSITE: PS00532; PEPC\_K\_ATP; 1.  
Glucose: Glucose; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
NP\_BIND: 233 240 ATP (BY SIMILARITY).  
CONFLICT 10 10 L > S (IN REF. 3).  
SEQUENCE 527 AA; 58300 MW; E55:EC80ZD15666 CRC64;

Query Match 8.8%; Score 87; DB 1; Length 527;  
Best Local Similarity 26.9%; Pred. No. 2.2;  
Matches 47; Conservative 24; Mismatches 70; Indels 34; Gaps 9;

QY 20 LNVGVKLLSAVEKMLV---OKGAPNGIEVVFGLLYALAAARTSPKVRASDVIFS 76  
DB 234 LSGTGKTT--SADADKLGCDDEHGSQDGVFNIEG-GCYAKCIHLSEK----EPQIFNA 288  
DY 77 NSFGE--RNVVVTGDLKKVLDGCAPLTFRTNKLRTFTTEAVVDFCIAYK-HKLPQL 133  
DB 289 TRFGVLENVVDDE-----TREAYDVSFTENTRA-----AYPIHMINI 330  
QY 134 NAAELGIPAEPSYLAADFLGTCPLKLSLQOSRKX--FASKYALK---TEGGVNV 133  
DB 331 VTPSMAGHPSAIVPLTADAFGLVPEISKLTKQVMYHFLSGVTSKLAGTERGVTS 385

RESULT 5  
YFUK\_ECOLI

ID YFUK\_ECOLI STANDARD; PRT; 729 AA.  
AC P52126;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein. yfjK.  
GN YFJK OR B2627.  
OS *Escherichia coli*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
Mau B., Shao Y.;  
-!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP)

RT "The complete genome sequence of *Escherichia coli* K-12.";  
RL Science 277:1453-1474(1997).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: U36840; AAA79736.1; -  
EMBL: AE008348; AAC75675.1; -  
PIR: T08639; T08639.  
EcoGene: EGI3197; yfjK.  
InterPro: IPR001410; DEAD.  
DR InterPro: IPR001650; Helicase\_C.  
DR Pfam: P500271; helicase\_C; 1.  
DR SMART: SM00487; DEXDC; 1.  
DR SMART: SM00490; HELIC\_C; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 729 AA; 83061 MW; CB65BCD1B36FBFAF CRC64;

Query Match 8.6%; Score 85.5; DB 1; Length 729;  
Best Local Similarity 22.4%; Pred. No. 4.4;  
Matches 41; Conservative 31; Mismatches 64; Indels 47; Gaps 8;

QY 2 ELKSDSNLSNLSVI--TDASSLNGVDKLLSAVEKMLVKGNAPNGIEVVFGLLYALAA 59  
DB 520 EMDNEPSSLVFHWTPATNFKLTKFAIIARLVPHPTFSRNGVEVKPTDWMIAKLAGYLSA 579  
QY 60 RTTSPKVRASDVIFS--SPGERNV--VVTGDLKKVLDGCAPLTFRTNKLRTFT 114  
DB 580 ESYSEYK---NQIDYAROWISEGEKRTLSIANDLKLI-----TN---TFGYT 623  
QY 115 FTEA-----YVDFCIAYK--HKLPOLNAAELGIPAEPSYLA 150  
DB 624 LPKVLNEDVYGHVHAKVGRGIRSKVDYTHVKLAFSEFHLPPGVNALEEIGIPQTLHRLV 683  
QY 151 DFL 153  
DB 684 DLL 686

RESULT 6  
PEPC\_RHOPA

ID PEPC\_RHOPA STANDARD; PRT; 537 AA.  
AC Q9ZKH4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 26-FEB-2003 (Rel. 41, Last annotation update)  
DE Phosphoenolpyruvate carboxykinase (ATP) (EC 4.1.1.49) (PEP carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
GN pckA.  
OS Rhodospseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Rhodospseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]  
RC SEQUENCE FROM N.A., SEQUENCE OF 1-12, AND DEVELOPMENTAL STAGE.  
RP STRAIN=7;  
RX MEDLINE=99235744; PubMed=10217755;  
RA Imai M., Nakata K., Roh J.H., Zahn K., Yukawa H.;  
PT "Molecular and functional characterization of the Rhodospseudomonas palustris no. 7 phosphoenolpyruvate carboxykinase gene.";  
RL J. Bacteriol. 181:2689-2696(1999).  
CC -!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate + CO(2).  
CC -!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING LOG PHASE WITH 10-20 FOLD REDUCTION AT ONSET OF STATIONARY PHASE.  
CC -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP)

CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EXBL; AB015618; BAA34956.1; -  
 CC DR HSSP; P22259; 1AQ2.  
 CC DR HAMAP; MF 00453; -; 1.  
 CC DR InterPro; IPR001272; PEPCK\_ATP.  
 CC DR Pfam; PF01293; PEPCK\_ATP\_1.  
 CC DR ProDom; PD004723; PEPCK\_ATP; 1.  
 CC DR TIGRFAMs; TIGR00224; pckA; 1.  
 CC DR PROSITE; PS03532; PEPCK\_ATP; FALSE NEG.  
 CC KW Glucocorticoid synthase; Lyase; Decarboxylase; ATP-binding.  
 CC FT NP B:NC 236 243 ATP (BY SIMILARITY).  
 CC ST SEQUENCE 537 AA; 58773 MW; 98F0ECICEDIC9EA8 CRC64;  
 CC  
 CC Query Match 8.5%; Score 84; DB 1; Length 537;  
 CC Best Local Similarity 26.7%; Pred. No. 4.2;  
 CC Matches 48; Conservative 21; Mismatches 73; Indels 38; Gaps 10;  
 CC  
 CC QY 20 LNVGDKKLLSAGEVKMLVOKGAPNEGIEVVFGLL--LYALAAPTTS---PKVCRADSDVI 74  
 CC Db 237 LSTGKTKLSADPNRTLIGDEHGKGGVFNTEGGCYAKCIKLSAENPEIYAA----- 291  
 CC QY 75 FSNISGGE--BNVVVTEGDLKKVLDGCGAPLRTFTNKLRTFGRTITEAYVDFCAAYK-HKLP 131  
 CC Db 292 -STRFGAVLEWVVLGDIKRPDPDGSK---TENTRS-----AYP-ESIP 332  
 CC QY 132 QLNAAELGTPADSVLAADFLGTCPLKSLQSRKY--FASWYALK---TEGVVN-TP 165  
 CC Db 333 NASLTGRAGQPKVWMLAADAFGVNPEIAKLTPAQAWYHFLSGYTKAVAGTERGVTEPT 392  
 CC  
 CC RESULT 7  
 CC TLDD\_BUCAI STANDARD; PRT; 483 AA.  
 CC AC P57478;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE TLDD protein homolog.  
 CC GN TLDD OR B0398.  
 CC OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 CC symbiotic bacterium).  
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC OC Enterobacteriaceae; Buchnera.  
 CC OX NCBI\_TaxID=118099;  
 CC [!]  
 CC RN SEQUENCE FROM N.A.  
 CC RP STRAIN=Tokyo 1998;  
 CC RC MEDLINE=20445173; PubMed=10993077;  
 CC RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 CC "Genome sequence of the endocellular bacterial symbiont of aphids  
 CC Buchnera sp. APS.";  
 CC RT Nature 407:81-86(2000).  
 CC RL  
 CC CC -!- SIMILARITY: BELONGS TO THE TLDD/PMBA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EXBL; AP001119; BAB13101.1; -  
 CC DR InterPro; IPR002510; PmbA\_TLDD.  
 CC

DR Pfam; PF01523; PmbA\_TLDD; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 483 AA; 52977 MW; A05CE98518720EBA CRC64;  
 Query Match 8.4%; Score 83.5; DB 1; Length 483;  
 Best Local Similarity 20.5%; Pred. No. 4.1;  
 Matches 32; Conservative 26; Mismatches 53; Indels 45; Gaps 5;  
 QY 65 KVRADSDVIFSNFGRNVVTEGDLKKVLD-----GCAPLTR 103  
 Db 305 KNORGSLIDDEGTPGKNILIIENGILKYNQDKLNARLMGVKSTGNGRRSYSLMPR 364  
 QY 104 FTKLRTFGRTTEAYVDFCIAYKHLPQLNAAELGIPAEKDSYLAADFGLGTCPLKSELQ 163  
 Db 365 MTN-----TYMLSGASKLDDIKSVDYGI-----YAVNFSG---GVQIT 401  
 QY 164 QSRKMFASMYALKTEGGVNVTPVSNLQGR-REVM 198  
 Db 402 SGKTFVFSSEAYLIKNGKVTPIKNTTLIGSGLEWM 437  
 CC  
 CC RESULT 8  
 CC ARGC\_LEPIN STANDARD; PRT; 385 AA.  
 CC ID ARGC\_LEPIN  
 CC AC Q8EYV8;  
 CC DT 15-SEP-2003 (Rel. 42, Created)  
 CC DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 CC DE N-acetyltransferase (EC 2.3.1.35) [Ornithine acetyltransferase;  
 CC Arginine biosynthesis bifunctional protein argJ includes: Glutamate  
 CC (Ornithine transacetylase) (OATase); Amino-acid acetyltransferase  
 CC (EC 2.3.1.1) (N-acetylglutamate synthase) (AGS)] [Contains: Arginine  
 CC biosynthesis bifunctional protein argJ alpha chain; Arginine  
 CC biosynthesis bifunctional protein argJ beta chain].  
 CC GN ARGJ OR LA4105.  
 CC OS Leptospira interrogans.  
 CC OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 CC OX NCBI\_TaxID=1173;  
 CC [!]  
 CC RN SEQUENCE FROM N.A.  
 CC RP STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 CC RX MEDLINE=22598143; PubMed=12712204;  
 CC RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,  
 CC Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,  
 CC Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,  
 CC Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,  
 CC Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Panchin A.,  
 CC Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,  
 CC Xu J.-G., Zhao G.-P.;  
 CC "Unique physiological and pathogenic features of Leptospira  
 CC interrogans revealed by whole-genome sequencing.";  
 CC RJ Nature 422:888-893(2003).  
 CC -!- FUNCTION: Catalyzes two activities which are involved in the  
 CC cyclic version of arginine biosynthesis: the synthesis of  
 CC acetylglutamate from glutamate and acetyl-CoA, and of ornithine by  
 CC transacetylation between acetylornithine and glutamate (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate = L-  
 CC ornithine + N-acetyl-L-glutamate.  
 CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + L-glutamate = CoA + N-acetyl-L-  
 CC glutamate.  
 CC -!- PATHWAY: Arginine biosynthesis; first step.  
 CC -!- PATHWAY: Arginine biosynthesis; fifth step.  
 CC -!- SUBUNIT: Heterotetramer of two alpha and two beta chains (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- MISCELLANEOUS: Some bacteria possess a monofunctional argJ, i.e.,  
 CC capable of catalyzing only the fifth step of the arginine  
 CC biosynthetic pathway.  
 CC -!- SIMILARITY: Belongs to the argJ family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

C the European Bioinformatics Institute. There are no restrictions on its  
C use by non-profit institutions as long as its content is in no way  
C modified and this statement is not removed. Usage by and for commercial  
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

R EMBL; AE011564; AAN51303.1; -  
R HAMAP; MF\_01106; -; 1.  
R InterPro; IPR002813; ArgJ.  
R Pfam; PF01960; ArgJ; 1.  
R ProDom; PD004193; ArgJ; 1.  
R TIGRFAMs; TIGR00120; ArgJ; 1.  
R Arginine biosynthesis; Multifunctional enzyme; Transferase;  
R Acyltransferase; Complete proteome.  
T CHAIN 1 178  
T ARGININE BIOSYNTHESIS BIFUNCTIONAL  
T PROTEIN ARGJ ALPHA CHAIN (BY SIMILARITY).  
T ARGININE BIOSYNTHESIS BIFUNCTIONAL  
T PROTEIN ARGJ BETA CHAIN (BY SIMILARITY).  
T CJEAVASE (NONHYDROLYTIC) (BY SIMILARITY).  
Q SEQUENCE 385 AA; 41445 MW; 5A523296D0AD7C2F CRC64;

Query Match 8.4%; Score 83; DB 1; Length 385;  
Best Local Similarity 21.3%; Pred. No. 3.4;  
Matches 29; Conservative 23; Mismatches 44; Indels 40; Gaps 4;  
Y 7 SNLSNLTIDASS---LNGVDKLLSAEVEKMLVQKAPN-----EG 45  
b 262 TKLEJTSIGASQAARKIGKISILNSPLVKRTIYGGDPNMGRLMAVGKVFDEPIPFEG 321  
Y 46 IEVFGILLALAAARTSPKQCRADSDVIFSNFGEENVVVTEGLAKVLGDCAPJTRFT 105  
b 322 LQIVFG---TLPKVEANPEFLKSLVKNNTETISLNVLVNLTGISMKPGC----- 373  
Y 106 NKLRTFGRFTFEAYVD 121  
b 371 -----DFTKVI 378

RESULT 9  
PCK AGST5 STANDARD; PRT; 536 AA.  
C Q80U94;  
T 28-FEB-2003 (Rel. 41, Created)  
T 28-FEB-2003 (Rel. 41, Last sequence update)  
T 28-FEB-2003 (Rel. 41, Last annotation update)  
E Phosphoenolpyruvate carboxylase [ATP] (EC 4.1.1.49) (PEP  
E carboxylase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
N PCKA OR AT00035 OR A3R\_C\_56.  
S Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
C Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
C Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
X NCBI\_TaxID=176299;  
P SEQUENCE FROM N.A.  
X MEDLINE=21608550; PubMed=11743193;  
A Wood D.W., Secubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
A Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
A Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
A Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C.,  
A Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
A Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon J.,  
A Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
A Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
A Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
A Nester E.W.;  
T "The genome of the natural genetic engineer Agrobacterium tumefaciens  
T C58.";  
J Science 294:2317-2323(2001).  
Y (2)

SEQUENCE FROM N.A.  
X MEDLINE=21608551; PubMed=11743194;  
A Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
A Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,

RA Houtiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
RA Wolam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,  
RA Cielo C., Slater S.;  
RT "Genome sequence of the plant pathogen and biotechnology agent  
RT Agrobacterium tumefaciens C58.";  
RL Science 294:2323-2328(2001).  
CC -!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate  
CC + CO(2).  
CC -!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxylase [ATP]  
CC family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE008978; AAL41066.1; ALT\_INIT.  
DR EMBL; AE007946; AAK85859.1; ALT\_INIT.  
DR HAMAP; MF\_00453; -; 1.  
DR InterPro; IPR001272; PEPCK\_ATP.  
DR Pfam; PF01293; PEPCK\_ATP; 1.  
DR ProDom; PD004723; PEPCK\_ATP; 1.  
DR TIGRFAMs; TIGR00224; PCKA; 1.  
DR PROSITE; PS00532; PEPCK\_ATP; 1.  
KW Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
FT NP\_BIND 236 243 ATP (BY SIMILARITY).  
SQ SEQUENCE 536 AA; 57882 MW; 6F90B8F484C86A5A CRC64;

Query Match 8.4%; Score 83; DB 1; Length 536;  
Best Local Similarity 26.8%; Pred. No. 5.1;  
Matches 47; Conservative 18; Mismatches 72; Indels 38; Gaps 9;

QY 20 LNVGVKKLLSAEVEKMLVQKAPNEGIEVVVFGLL--LYALAAARTTS---PKVQRADSDV 74  
DB 237 LSGTGKTTLSADPARTLIGDDEHGWGHEGIFNFGGCGYAKAKLSSEAPEPIYAA----- 291  
QY 75 FSNFGE--RNVVVTEGLKKVLDGCGAPLRTFRNKLRTFTTEAYVDFCIAYK-HKLP 131  
DB 292 -TNRFGTGVLENVLDSESRVDFNDN--SLTENTRS-----AYLHPFP 331  
QY 132 CLNKAARELGIPAEYSVAADFLGTCPKSEL--QOSRKMFAWMYALK---TEGV 181  
DB 332 NASETGIAGHPKT-IMLTADAFGLVLPPIARLTPEQAWYHFLSGYTKAVAGTEKGV 386

RESULT 10  
ACKA\_THETN STANDARD; PRT; 401 AA.  
ID ACKA\_THETN  
AC Q889V4;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Acetate kinase [EC 2.7.2.1] (Acetokinae).  
GN ACKA OR TTE1481.  
CS Thermoanaerobacter tengcongensis.  
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
CC Thermoanaerobacteriaceae; Thermoanaerobacter.  
OX NCBI\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4 / JCM 11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Lu Y., Lai X., Huang J., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
RT "A complete sequence of T. tengcongensis genome.";  
RL Genome Res. 12:689-700(2002).

```

-!- CATALYTIC ACTIVITY: ATP + acetate = ADP + acetyl phosphate.
-!- PATHWAY: Conversion of acetate to acetyl-CoA: first step.
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-!- SIMILARITY: Belongs to the aceto kinase family.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE013105; RAN24700.1; -.
HAMAP; MF_00020; -.
InterPro; IPR000896; Acetate_kin.
InterPro; IPR004372; ACKA.
Pfam; PF00871; Acetate_kinase; 1.
TIGRFAMs; TIGR00016; ACKA; -.
PROSITE; PS01075; ACETATE_KINASE_1; 1.
PROSITE; PS01076; ACETATE_KINASE_2; 1.
Transferrase; Kinase; Complete proteome.
SEQUENCE 401 AA; 44361 MW; FD71C642F1E019B0 CRC64;

Query Match      8.2%; Score 81.5; DB 1; Length 401;
Best Local Similarity 19.3%; Pred. No. 5;
Matches 40; Conservative 35; Mismatches 87; Indels 45; Gaps 7;

22 GVDKLLGAEV---FMVLVQKGFN--EGIEVVGGLL-YALAAATTPKVGQADSDVIF 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 GINDSLTHQAEGKEKVKIQREMKHKEAIQVLEVLVDKEIGV-KMKKEIDAVGHRVWH 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

76 SNSFGERNVWVTEGDLKKVLQSC---APLTRTN-----KLTFTFTIE 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

95 GGEYTFDSVLIDDEVIRK-JEDCIDLAPLNANIEG-KAQQIMPGVPMVAVFETAFHQ 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

118 AYVDFCIAY-----KHKLPQNAAAELGIPAEISYLAADPLG-GCP 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

154 TYPDYAYIYFIEYVEYKXRIERYGHGTSHKYSVMRAAILGRPIELKIVTCHJNGA 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

158 KLSLEQQRKMPASMYALKTEGVNVT 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

214 SITAVKNGKSIDTSMGFTPLEGLAVGT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

SULT 11
N2_BORBU
LON2_BORBU STANDARD; PRT; 813 AA.
OS1558;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ATP-dependent protease Ia homology (EC 3.4.21.-).
BB0613.
Borrelia burgdorferi (Lyme disease spirochete).
Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
NCBI_TaxID=139;
[1]
SEQUENCE FROM N.A.
STRAIN=ATCC 35210 / B31;
MEDLINE=98065943; PubMed=9403685;
Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
Peterson J., Kervlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
Van Vugt R., Palmer N., Adams M.D., Gocayne J.C., Weidman J.,
Utterback T., Wathley L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
"Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi."
Nature 390:580-586(1997).
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S16.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE031162; AAC66962.1; -.
CC PIR; D70176; D70176.
CC MEROPS; S16.UPW; -.
CC TIGR; BB6613; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_cent.
CC InterPro; IPR003111; LON.
CC InterPro; IPR001984; Lon_endopep.
CC InterPro; IPR004815; Lon_fam.
CC Pfam; PF00004; AAA; 1.
CC Pfam; PF02190; LON; 1.
CC Pfam; PF05362; Lon_C; 1.
CC PRINTS; PR00830; ENDOLAPTASE.
CC SMART; SM00382; AAA; 1.
CC SMART; SM00464; LON; 1.
CC TIGRFAMs; TIGR00763; Lon; 1.
CC PROSITE; PS01046; LON_SER; 1.
KW Hydrolase; Serine protease; ATP-binding; Complete proteome.
FT NP_BOND 369 376 ATP (POTENTIAL).
FT ACT_SITE 719 719 BY SIMILARITY.
SQ SEQUENCE 813 AA; 92312 MW; 5129AA1498C5D0F6 CRC64;

Query Match      8.2%; Score 81.5; DB 1; Length 813;
Best Local Similarity 31.3%; Pred. No. 12; Indels 17; Gaps 6;
Matches 37; Conservative 19; Mismatches 43;

Cy 8 NLSN-LVITDASSNGVDKLLSAEVEKMLVQKGFNIEGVVVGGLL-YALAAATTPSKV 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 DLSNVLVFTVITANSNGMSKPLLD-RYEIIKVGYSYIEKLEIAKIFLPSIKESFLDKV 533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Cy 67 Q-RADSDVIFSNFGRNVVVTG--DLKKVLDGCAPLTRFTNKLRTFTFTTEAY 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 534 YRIEDDVIFNLI---RNYTMESGVRGLKRVL-----TNLIRRLVRELLYEV 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
SPSY_MOUSE
ID SPSY_MOUSE STANDARD; PRT; 366 AA.
AC P97355;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Spermine synthase (EC 2.5.1.22) (Spermidine aminopropyltransferase)
DE (SPMSY).
GN SMS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strom T.M., Francis F., Lorenz B., Boeddrich A., Econs M.J.,
RA Lehrach H., Weittinger T.;
RA "Pex gene deletions in Gy and Hyp mice provide mouse models for
RA X-linked hypophosphatemia."
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Diaphragm;
RC Niiranen K., Korhonen V., Janne J.;
RA "Nucleotide sequence of mouse spermidine aminopropyltransferase
RT cDNA."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

```

```

SEQUENCE OF 316-366 FROM N.A.
MEDLINE=98133937; PubMed=94673015;
Lorenz B., Francis F., Gempel K., Boeddrich A., Joster M., Schmahl W.,
Schmidt J., Lehrach H., Meitinger T., Strom T.M.;
"Spermine deficiency in Gy mice caused by deletion of the spermine
synthase gene.";
Hum. Mol. Genet. 7:541-547(1998).
-!- CATALYTIC ACTIVITY: S-adenosylmethioninamide + spermidine = 5'-
methylthioadenosine + spermine.
-!- PATHWAY: BIOSYNTHESIS OF SPERMINE FROM SPERMIDINE.
-!- SIMILARITY: BELONGS TO THE SPERMINE/SPERMINE SYNTHASES FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; Y09419; CAA70573.1;
EMBL; A5031486; AAB86631.1;
EMBL; AJ000093; CAA03919.1;
EMBL; AJ000087; CAA03918.1;
EMBL; AJ000088; CAA03918.1; JOINED.
EMBL; AJ000089; CAA03918.1; JOINED.
EMBL; AJ000090; CAA03918.1; JOINED.
EMBL; AJ000091; CAA03918.1; JOINED.
EMBL; AJ000092; CAA03919.1; JOINED.
MGD; MGI:109490; Sns.
GO; GO:0008215; P: spermine metabolism; IMP.
InterPro: IPR001045; Spermine synthase.
Pfam; PF01564; Spermine synth; 1.
PROSITE; PS01330; SPERMIDINE SYNTHASE; 1.
Transferase.
DOMAIN 192 232 BINDING TO DECARBOXYLATED SAM
(POTENTIAL).
SEQUENCE 366 AA; 41313 MW; D549F319F31C43C5 CRC64;
Query Match 9.1%; Score 80; DB 1; Length 366;
Best Local Similarity 22.9%; Pred. No. 6.1;
Matches 52; Conservative 34; Mismatches 65; Indels 76; Gaps 13;
24 DKLLSAEVEKMLVQKARNEGIEV-----FGLLYALAARTTSKVRQADSDVFSNSF 79
128 DGRLEVEYDIEVVYDEDSYPQNIKLRKQFGNIL-----ILSGVNLAEISLAYTTRAI 181
80 -----GERNVVTEGP-----LK-----KVLGGCAPLTFRTNKLRT 110
182 MGGGKEGYTKDVLILGGGGGILCEIVLKKPKVMTWVIDQWIDGGCKKYR-----RT 236
111 FGRTFT-----EAYVDFCI-----AYKHKLPLQNAAEJGIPAEDSYLAADF 152
237 CGDVLNLRGCGYQVLIEDCIPVWKYAKEGREFDVINDLTAVPISLTSPEEDS--TWDF 294
153 LGTCPKLS--ELQSRKRFASMTALKEGVVN--TPVSNL--RQJGR 194
295 LRJLDSLMKVLKQDGKYF-----TOGNCVNLTEALSLEYEQJGR 334

```

SULT 13

CA\_SYNY3

OTCA\_SYNY3 STANDARD; PRT; 308 AA.

Q55497;

15-JUL-1998 (Rel. 36, Created;

15-JUL-1998 (Rel. 36, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)

Ornithine carbamoyltransferase, anabolic (EC 2.2.3.3) (OTCase).

AKRF OR SLL0902.

Synecocystis sp. (strain PCC 6803).

Bacteria; Cyanobacteria; Chroococcales; Synecocystis.

NCBI\_TaxID=1148;

[1]

```

SEQUENCE FROM N.A.
MEDLINE=96127529; PubMed=9590279;
Kareko T., Tanaka A., Sato S., Kotani H., Suzuki T., Miyajima N.,
Sugitara M., Tabata S.;
"Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
region from map positions 64% to 92% of the genome.";
J. Mol. Biol. 253:153-166(1995).
-!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
+ L-citrulline.
-!- PATHWAY: Arginine biosynthesis.
-!- SUBCELLULAR LOCATION: Cytoplasmic.
-!- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; D64036; BAA10847.1;
EMBL; PIR; S76000; S76000.
DR HSSP; P04391; LAKM.
DR InterPro; IPR006130; Asp/Orn Cotransf.
DR InterPro; IPR002292; Orn carbTransf.
DR InterPro; IPR006131; OTCase_O.
DR InterPro; IPR006132; OTCase_P.
DR Pfam; PF0185; OTCase; 1.
DR Pfam; PF02729; OTCase N; 1.
DR PRINTS; PR00100; AOTCASE.
DR TIGRFAMs; TIGR00658; orn carb tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
Arginine biosynthesis; Transferase; Complete proteome.
KW SEQUENCE 308 AA; 33616 MW; C6CD2E4998592CFA CRC64;
Query Match 8.0%; Score 79; DB 1; Length 308;
Best Local Similarity 24.8%; Pred. No. 6.1;
Matches 39; Conservative 25; Mismatches 63; Indels 30; Gaps 7;
CY 28 LSAEVEKMLVQKAPNEGIEV-----EVVFGLLLYALAARTTSKVRQADSDVFPNSFGE 81
DB 17 LTTEEMKSLQLLAADKSGVLKPHCKKILGLLFYKASTRT---RVSFTAAWYQLGGQVLD 73
QY 82 RNVVYTE-GDLKKVLDGCGAPLTFRTNKL--RTFGRTFEAYVDFCIAYKHKLPLQNAAAE 138
DB 74 LNPSTVQVGRGEPIQDARVLDYIDILAVRTFKCDLQTFADHA-----KYPINALSD 126
QY 139 LGIPAEEDSYLAADFCTCPKLSLQSRKRFASMTALKEGVVN--TPVSNL--RQJGR 175
DB 129 LEHP-----COILADLQTIKECFGKJGL 152
RESULT 14
CPXC_AGRU STANDARD; PRT; 422 AA.
AC P24466;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450-p1aF1, plant-inducible (EC 1.14.-.-).
GN CYP103 OR P1NF1 OR VIRH1.
OS Agrobacterium tumefaciens.
OC P-aseid pt1A6.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OC NCBI_TaxID=358;
[1]
RP SEQUENCE FROM N.A.
MEDLINE=89213933; PubMed=2708311;
Kanamoto R.H., Powell A.T., Akiyoshi D.E., Regier D.A.,
Kerstetter R.A., Nester E.W., Hawes M.C., Gordon M.P.;

```

```

"nucleotide sequence and analysis of the plant-inducible locus pInF
from Agrobacterium tumefaciens";
J. Bacteriol. 171:2506-2512(1989).
-!- FUNCTION: NOT ESSENTIAL FOR VIRULENCE, BUT MAY BE INVOLVED IN THE
DETOXICATION OF PLANT PROTECTIVE AGENTS AT THE SITE OF WOUNDING.
-!- INDUCTION: TRANSCRIPTIONALLY ACTIVATED IN THE PRESENCE OF WOUNDED
PLANT TISSUE AND BY PLANT PHENOLIC COMPOUNDS, SUCH AS
ACETOSYRINGONE
-!- SIMILARITY: Belongs to the cytochrome P450 family.
-----
This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M19352; AAA82502.1; .
PIR; A32306; A32306.
InterPro; IPR001128; Cytochrome_P450.
Pfam; PF00067; P450; 1.
PRINTS; PR00385; P450.
PROSITE; PS00086; CYTOCHROME_P450; 1.
Oxidoreductase; Monooxygenase; Electron transport; Heme; Plasmid.
METAL 369 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SEQUENCE 422 AA; 47519 MW; 6A9PE4AA95E2302 CRC64;
Query Match 8.0%; Score 79; DB 1; Length 422;
Best Local Similarity 23.6%; Pred. No. 8.9;
Matches 51; Conservative 29; Mismatches 60; Indels 76; Gaps 14;
20 LAGVDKLLSAEVEKMLVQKGFAPNEGIE--VYFGLLLYALAARTTSPKVQRADSDVIPS 77
70 LGTDPRTCQICETELM-----NRGKAGAVDFI-----DHSMLFSN 107
78 --SFGERNVYVTEGDLKYLDCAP-LTRFT-----NKLRTFGRTTEAYVDFCIAYX 127
108 GETHGKRRLGSLKAFSFRVVEALRPEIAKITECLWDLQKVDFF--NTIEMVAS----- 159
128 HKLPQLNAAAEIGIPAE-----SYLAADF-GICPRLSEIQ----- 163
160 -QLPALTIASVLGLPSEDTPFFTRLVYKSRCLSPSWRDEFEETIEASAIQLQDVRSVI 218
164 --QSRKM---FASMYALKT--EGGVNTPVSNLRQ- 192
219 ADSGRMRDQFLSRY-JKAVREAGTL-SPIEIMQL 252
-----
RESULT 15
PCK_FUSNN STANDARD; PRT; 527 AA.
QOREI2;
28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP
carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
PCKA OR FN1120.
Fusobacterium nucleatum (subsp. nucleatum).
Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
Fusobacterium.
NCBI_TaxID=76856;
[1]
SEQUENCE FROM N. A.
STRAIN=ATCC 25586;
MEDLINE=21866394; PubMed=11889.09;
Kapatal V., Andersson I., Ivanova N., Reznik G., Los T., Lykidis A.,
Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
Fonstein X., Kyrpides N., Overbeek R.;
"Genome sequence and analysis of the oral bacterium Fusobacterium

```

```

nucleatum strain ATCC 25586 ";
J. Bacteriol. 184:2005-2018(2002).
-!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate
+ CO(2) .
-!- PATHWAY: Rate-limiting cytochrome P450 enzyme.
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [ATP]
family.
-----
This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AE010616; AAL95316.1; .
HAMAP; MF_00453; .
InterPro; IPR001272; PEPCK_ATP.
Pfam; PF01293; PEPCK_ATP; 1.
PRODOM; PD004723; PEPCK_ATP; 1.
TIGRFAMs; TIGR00224; pckA; 1.
PROSITE; PS00532; PEPCK_ATP; 1.
KW Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.
FT NP_BIND 230 237 ATP (BY SIMILARITY).
SQ SEQUENCE 527 AA; 59055 MW; 275849FDF254AC01 CRC64;
Query Match 8.0%; Score 79; DB 1; Length 527;
Best Local Similarity 26.6%; Pred. No. 12;
Matches 46; Conservative 26; Mismatches 67; Indels 34; Gaps 10;
QY 20 LAGVDKLLSAEVEKMLV---QKGFAPNEGIEVYFGLLLYALAARTTSPKVQRADSDVIPS 76
Db 231 LSGTGRKTLSDPNRKLGDDEHGWCDKGIHFNEG---GCYAKCINLK-ESESEPEIYRA 285
QY 77 NSFGE--RNWVTEGGLKKVLCGAPLRTFNKLTFRGRTT-EAYVDFCIAYKHLPLQL 133
Db 286 IKGSLVENVVD-----PITR---KIYEDASITENTRUGPIDY---TPNA 327
QY 134 NAAAEELG:PAEDSYLAADFICGTCPLSELQSSRM--FASMYALK---TEGGV 181
Db 328 ELSGVGGIPKWIIFLTADSRGLPPIPSRLSLQEAAMVHFVTGFTAKLAGTELGV 380
Search completed: November 7, 2003, 14:49:01
Job time : 18 secs

```

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 CompuGen Ltd.

Query protein - protein search, using sw model

Run on: November 7, 2003, 14:46:13 : Search time 34 Seconds  
 (without alignments):  
 1502.777 Million cell updates/sec

File: US-09-613-486-15  
 Effect score: 991  
 Sequence: 1 MELMSDSNLSNLVITDASSL.....GGVNVTPVSNLRQJGRREVW 198

Scoring table: BLCSJM62  
 Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database: SPTREMBL23.\*

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.podent.\*
- 12: sp.virus.\*
- 13: sp.vertibrate.\*
- 14: sp.unclassified.\*
- 15: sp.virus.\*
- 16: sp.bacteriap.\*
- 17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

sult No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	198	12	071213 grapevine 1
2	984	99.3	198	12	039857 grapevine 1
3	921	92.9	198	12	038223 grapevine 1
4	297	30.0	204	12	090629 sugar beet
5	296	29.9	204	12	096664 sugar beet
6	295	29.8	204	12	089910 sugar beet
7	293	29.6	204	12	031724 sugar beet
8	286	28.9	204	12	008531 sugar beet
9	266.5	26.9	215	12	085858 beet yellow
10	130.5	13.2	223	12	080711 citrus tris
11	129.5	13.1	223	12	090EG4 citrus tris
12	125.5	12.7	222	12	0930P0 citrus tris
13	125.5	12.7	223	12	0930P2 citrus tris
14	125.5	12.7	223	12	090EG5 citrus tris
15	125	12.6	223	12	083T3 citrus tris
16	124.5	12.6	223	12	090EG2 citrus tris

17	123.5	12.5	223	12	09E7M2 citrus tris
18	123.5	12.5	223	12	09E7M4 citrus tris
19	123.5	12.5	223	12	09PXT0 citrus tris
20	123.5	12.5	223	12	09QEG3 citrus tris
21	123	12.4	212	12	09DWM1 citrus tris
22	123	12.4	223	12	09E7M3 citrus tris
23	120.5	12.2	195	12	08V1P7 citrus tris
24	120.5	12.2	222	12	090DN9 citrus tris
25	120.5	12.2	223	12	09QEG1 citrus tris
26	120.5	12.2	223	12	089948 citrus tris
27	120.5	12.2	223	12	0990P1 citrus tris
28	120.5	12.2	223	12	09WID9 citrus tris
29	120.5	12.2	223	12	091FW4 citrus tris
30	120.5	12.2	223	12	006192 citrus tris
31	120.5	12.2	223	12	09PXT3 citrus tris
32	118.5	12.0	223	12	08JYU9 citrus tris
33	117.5	11.9	223	12	09PXS9 citrus tris
34	117.5	11.9	223	12	0991G6 citrus tris
35	116	11.7	223	12	09QEG6 citrus tris
36	115.5	11.7	195	12	08V1P5 citrus tris
37	115.5	11.7	223	12	08JYU0 citrus tris
38	115	11.6	223	12	09PXT1 citrus tris
39	114.5	11.6	223	12	09D7F8 citrus tris
40	114.5	11.6	223	12	010474 citrus tris
41	114.5	11.6	223	12	08B3R9 citrus tris
42	113.5	11.5	194	12	08V1P6 citrus tris
43	111.5	11.3	223	12	08JYU8 citrus tris
44	110.5	11.2	223	12	09PXT2 citrus tris
45	109	11.0	203	12	099AT6 citrus tris

## ALIGNMENTS

## RESULT 1

071213 PRELIMINARY: PRT; 198 AA.  
 AC 071213;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DC 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 D7 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE 22 kDa coat protein.  
 OS Grapevine leafroll-associated virus 2.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
 OC Closterovirus.  
 OX NCBI\_TaxID=64003;  
 RN [1].  
 RX SEQUENCE FROM N.A.  
 RP MEDLINE=98264507; PubMed=9603345;  
 RA Zhu H.Y., Ling K.S., Gosczyrski J.E., McPerson J.R., Gonsalves D.;  
 RT "Nucleotide Sequence and Genome Organization of Grapevine Leafroll-  
 RT Associated Virus-2 are Similar to Beet Yellows Virus, the  
 RT Closterovirus Type Member.";  
 RL J. Gen. Virol. 79:1289-1298(1998).  
 DR EMBL; AF039204; AAC40861.1; -;  
 DR InterPro; IPR002679; Closter\_coat.  
 DR Pfam; PF01785; Closter\_coat; 1.  
 KW Coat protein.  
 SQ SEQUENCE 198 AA; 21660 MW; 8ED85EF2EFD1DFB03 CRC64;

Query Match 100.0%; Score 991; DB 12; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-89;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MELMSDSNLSNLVITDASS:NGVDKLLSAEVEKMLVQKGNEGTEVVFGLLVALAAR 60
Ds	1	MELMSDSNLSNLVITDASS:NGVDKLLSAEVEKMLVQKGNEGTEVVFGLLVALAAR 60
Qy	61	TTSPKQVQADSDVIFNSFGERNVVTGDLKKVLDGCAPLTFRTTNKLTFRGTFTEAYV 120
Ds	61	TTSPKQVQADSDVIFNSFGERNVVTGDLKKVLDGCAPLTFRTTNKLTFRGTFTEAYV 120
Qy	121	DFCIAYKHKLQPNAAALGIPAEDSYLAADFLGTCTCKLSLQQSRKMFASMYALKTEGG 190

121 DFCAYKHKLPOLNAAALG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
181 VVNTPVSNLRQLGRREV 198  
181 VVNTPVSNLRQLGRREV 198

## SULT 2

9857 O39857 PRELIMINARY; PRT; 198 AA.  
O39857;  
01-JAN-1998 (TReMBLrel. 05, Created)  
01-JAN-1998 (TReMBLrel. 05, Last sequence update);  
01-DEC-2001 (TReMBLrel. 19, Last annotation update);  
Coat protein.  
Grapevine leafroll-associated virus 2.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
Closterovirus.  
NCBI\_TaxID=64003;  
[1]  
SEQUENCE FROM N.A.  
Abou-Gharem N.;  
"The nucleotide sequence of the 3'terminal region of grapevine  
leafroll associated closterovirus 2.";  
Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
EMBL: Y14131; CAA74566.1; -;  
InterPro: IPR002679; Closter\_coat.  
Pfam: PF01785; Closter\_coat; 1.  
SEQUENCE 198 AA; 21630 MW; 15D631E4088D3F01 CRC64;

Query Match 99.3%; Score 984; DB 12; Length 198;  
Best Local Similarity 99.5%; Pred. No. 3.7e-88;  
Matches 197; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 MELMSDSNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAAR 60  
1 MELMSDSNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAAR 60  
61 TTSPKQVQADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
61 TTSPKQVQADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
121 DFCIAYKHKLPOLNAAALG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
121 DFCIAYKHKLPOLNAAALG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
181 VVNTPVSNLRQLGRREV 198  
181 VVNTPVSNLRQLGRREV 198

## SULT 3

9857 O39857 PRELIMINARY; PRT; 198 AA.  
O39857;  
01-MAR-2003 (TReMBLrel. 23, Created)  
01-MAR-2003 (TReMBLrel. 23, Last sequence update);  
01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
Coat protein.  
Grapevine rootstock stem lesion associated virus.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
Closterovirus.  
NCBI\_TaxID=167634;  
[1]  
SEQUENCE FROM N.A.  
Zhang Y., Rowhani A.;  
"Nucleotide sequence of grapevine rootstock stem lesion associated  
virus.";  
Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
EMBL: AF314061; AAN63472.1; -;  
SEQUENCE 198 AA; 21620 MW; 370E95A9369734A9 CRC64;

Query Match 92.9%; Score 921; DB 12; Length 198;  
Best Local Similarity 90.9%; Pred. No. 5.4e-82;  
Matches 180; Conservative 10; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MELMSDSNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAAR 60  
DB 1 MELMSDDNLSGLVITDASSLNGVDKLLSAEVEKMLVOKGAPSQGIEVVFGLLLYALAAR 60  
QY 61 TTSPKQVQADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
DB 61 TTSPKQVQADSDVIFSNFGERNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYV 120  
QY 121 DFCIAYKHKLPOLNAAALG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGG 180  
DB 121 DFCVAYKHKLPOLNAAALG:PAEDSYLAADFLGACPKLSELOQSRKVFASMYALKTEGG 180  
QY 181 VVNTPVSNLRQLGRREV 198  
DB 181 VVNTPVSNLRQLGRREV 198  
RESULT 4  
Q39629  
ID Q39629 PRELIMINARY; PRT; 204 AA.  
AC Q39629;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2003 (TReMBLrel. 13, Last annotation update);  
DE Major capsid protein.  
DE Sugar beet yellow virus (SBYV).  
CC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
CC Closterovirus.  
OX NCBI\_TaxID=12161;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BYV-4;  
RX MEDLINE=20079557; PubMed=10611288;  
RA Peremyslov V.V., Hagiwara Y., Dolja V.V.;  
RT "HSP70 homolog functions in cell-to-cell movement of a plant virus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:14771-14776(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BYV-4;  
RA Peremyslov V.V., Hagiwara Y., Alzhanova D., Dolja V.V.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF190581; AAF14305.1; -;  
DR InterPro: IPR002679; Closter\_coat.  
DR Pfam: PF31785; Closter\_coat; 1.  
SQ SEQUENCE 204 AA; 22419 MW; FD6337E1D2490D33 CRC64;

Query Match 30.0%; Score 297; DB 12; Length 204;  
Best Local Similarity 37.6%; Pred. No. 6.4e-21;  
Matches 71; Conservative 32; Mismatches 84; Indels 2; Gaps 2;  
QY 7 SNLSNLVITDASSLNGVDKLLSAEVEKMLVOKGAPNEGIEVVFGLLLYALAARTSPKV 66  
DB 11 ATFENVSLADQTCJMGEDCDKLRDFECLKLGVPEDKLGALGLCLYSCATIGTSNKV 70  
QY 67 QRAQSDVIFSNFGE-RNVVVTGGDKKVLGGCAPLRTFTNKLRTFGRTTEAYVDFCIA 125  
DB 71 SVQPTSTF:KASFGSGKELFLTHGELRSFLDSQKLEGGKPKKRCFCRTFKDYISFAKE 130  
QY 126 YXHKLPOLNAAALG:PAEDSYLAADFLGTCPKLSELOQSRKVFASMYALKTEGGVWNTP 185  
DB 131 YRGRUPPIARANPHGLPAEDHYLAADF:STSTELTDLQGRLLARENATHTFS-SESP 189  
QY 186 VSNLRQLGR 194  
DB 190 VTSLKQLGR 198

RESULT 5  
Q39664



```

SQ SEQUENCE 204 AA; 22446 MW; D913184ECC93883B CRC64;

Query Match 29.8%; Score 295; DB 12; Length 204;
Best Local Similarity 37.6%; Pred. No. 1e-20;
Matches 71; Conservative 31; Mismatches 85; Indels 2; Gaps 66

QY 7 SNLSNVLVITDASSLNGVDKLLSAEVEKMLVKQAPNEGIEVVFGLLLYALAAARTTSPKV 66
DB 11 ATPENVSLVDQTC LHGEDCDKLRNFECKLKGVPEDKLGALGLCLYSCATIGTGNKV 72
QY 67 QRADSDVIFSNFSF-GRNVVVTEGDLKKVLDGCAPLTRTNKLRFTGRFTTEAYVDFCIA 125
DB 71 SVQFTSTFTIKASFGGKKE-FLTHGELSRFLDSQKLEGRPNKLRFCRTCFKDYISFAKE 130
QY 126 YKHKLPMQNAAEELGIPAEDSYLAADFLGTCPKLSELOQRKMFASMYALKTEGGVNT 185
DB 131 YGRGLPPIARANPHGLUPAEHDYLAADFISTSTELTDLOQRULLARENATHTEFS-SESP 189
QY 186 VSNLRQLGR 194
DB 190 VTSLKQLGR 198

RESULT 7
Q03724 PRELIMINARY; PRT: 204 AA.
AC Q03724;
DT 01-NOV-1996 (TEMBLrel. C1, Created)
LT 01-NOV-1996 (TEMBLrel. C1, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. C3, Last annotation update)
DE Coar protein.
OS Sugar beet yellow virus (SBYV).
CC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
CC Closterovirus.
CX NCB:TaxID=12161;
RN [1]
RP SEQUENCE FROM N.A.
RC STRA-N=Broom's barn;
RA Brunstedt J., Moseley J.J., Hull R.;
RT "Nucleotide sequence of cDNA encoding the coat protein of beet yellows virus.";
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=German;
RA Agraciovsky A., Koonin E.Y., Boyko V.P., Maiss E., Lunina N.A.,
RA Akabekov J.G.;
RT "Beet yellows closterovirus: complete genome structure and
RT identification of a leader papain-like thiol protease.";
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: M59452; AAA72955.1; --
DR EMBL: X73475; CAA51861.1; --
DR InterPro: IPR002679; Closter_coat.
DR Pfam: PF01785; Closter_coat; 1.
SQ SEQUENCE 204 AA; 22388 MW; 92B2A4DCA77543D9 CRC64;

Query Match 29.6%; Score 293; DB 12; Length 204;
Best Local Similarity 37.6%; Pred. No. 1.6e-20;
Matches 71; Conservative 31; Mismatches 85; Indels 2; Gaps 66

QY 7 SNLSNVLVITDASSLNGVDKLLSAEVEKMLVKQAPNEGIEVVFGLLLYALAAARTTSPKV 66
DB 11 ATPENVSLVDQTC LHGEDCDKLRNFECKLKGVPEDKLGALGLCLYSCATIGTGNKV 70
QY 67 QRADSDVIFSNFSF-GRNVVVTEGDLKKVLDGCAPLTRTNKLRFTGRFTTEAYVDFCIA 125
DB 71 SVQFTSTFTIKASFGGKKE-FLTHGELSRFLDSQKLEGRPNKLRFCRTCFKDYISFAKE 130
QY 126 YKHKLPMQNAAEELGIPAEDSYLAADFLGTCPKLSELOQRKMFASMYALKTEGGVNT 185
DB 131 YGRGLPPIARANPHGLUPAEHDYLAADFISTSTELTDLOQRULLARENATHTEFS-SESP 189
QY 186 VSNLRQLGR 194

```





133 -----TNALRVGRTDIALYLAFC-RQNRNLSYGGRLPDAGIPAGVHYLCADFL- 180  
155 TCKPLSELOQSRKQFASMYALKTEGG--VWNTPVSNLRQLGR 194  
181 TGAGLTDECAVYIOAKQLKKGADDEVV---VTNVRQLGK 219

SUJT 14  
Q9QSGS PRELIMINARY; PRT: 223 AA.  
Q9QSGS;  
01-MAY-2000 (TrEMBLrel. 13, Created)  
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
Coat protein.  
Citrus tristeza virus.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
Closterovirus.  
NCBI\_TaxID=12162;  
(1)\_TaxID=12162;  
SEQUENCE FROM N.A.  
STRAIN=19-21;  
Bonacalza B., Febres V., Niblett C.L., Nolasco G.;  
"Biomolecular characterization of Citrus Tristeza Virus isolates from Portugal";  
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
(2)  
SEQUENCE FROM N.A.  
STRAIN=19-21;  
Sequeira Z., Nolasco G.;  
"Bacterial expressed coat protein for the development of a single antibody for routine detection of Citrus Tristeza Virus in Portugal";  
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
EMBL: AF184114; AAD56567.1;  
InterPro: IPR002679; Closter\_coat.  
Pfam: PF01785; Closter\_coat.1.  
SEQUENCE 223 AA; 24886 MW; 20723BBA:9E2CBFF CRC64;

Query Match 12.7%; Score 125.5; DB 12; Length 223;  
Best Local Similarity 27.0%; Pred. No. 0.00044;  
Matches 60; Conservative 27; Mismatches 74; Indels 61; Gaps 11;  
1 MELMSDSNLSNLVITDASLNGVDKLLSAEYKMLVQKG----APNEGIEVVFGLLLYA 56  
31 MNLHIDPTL--TAMNDVROLGTQQNALNRDL--FLTLKGKHPNLPKDKDFH-AMWLYR 86  
57 LAARTSPKVRAD-----SDVIF-SNSFGERNVVVTEGLKKV 94  
87 LAVKSS--LQSCDDTTGITYTREGVEVDLPKLTWDVFNKSGIGNR----- 132  
95 LDGCAPLFTFTNKLRTFGRTTEAYVDFCIAYKHKLPQLNAAAEGLGIPAEISYLAADFGLG 154  
133 -----TNALRVGRTDIALYLAFC-RQNRNLSYGGRLPDAGIPAGVHYLCADFL- 180  
155 TCKPLSELOQSRKQFASMYALKTEGG--VWNTPVSNLRQLGR 194  
191 TGAGLTDECAVYIOAKQLKKGADDEVV---VTNVRQLGK 219

SUJT 15  
Q8B3T3 PRELIMINARY; PRT: 223 AA.  
Q8B3T3;  
01-MAR-2003 (TrEMBLrel. 23, Created)  
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
25 kDa major coat protein.  
Citrus tristeza virus (isolate T36) (CTV).  
Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
Closterovirus.  
NCBI\_TaxID=31712;  
(1)

RP SEQUENCE FROM N.A.  
RX MEDLINE=99307427; PubMed=10377432;  
RA Satyanarayana T., Gowda S., Boyke V.P., Albiach-Marti M.R.,  
Mawassi M., Navas-Castillo J., Karasev A.V., Dolja V., Hill M.B.,  
Lewandowski D.J., Moreno P., Bar-Joseph M., Gartsey S.M., Dawson W.O.;  
"An engineered closterovirus RNA replicon and analysis of heterologous  
terminal sequences for replication";  
Proc. Natl. Acad. Sci. U.S.A. 96:7433-7438(1999).  
RN (2)  
RP SEQUENCE FROM N.A.  
RA Satyanarayana T., Gowda S., Ayllon M.A., Dawson W.C.;  
"Frameshift mutations in infectious cDNA clones of Citrus tristeza  
virus: A strategy to minimize the toxicity of viral sequences to  
Escherichia coli";  
Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY170468; AAO12723.1;  
KW Coat protein.  
SQ SEQUENCE 223 AA; 24890 MW; 5132F5E753023061 CRC64;  
Query Match 12.6%; Score 125; DB 12; Length 223;  
Best Local Similarity 25.9%; Pred. No. 0.00049;  
Matches 56; Conservative 33; Mismatches 63; Indels 64; Gaps 10;  
QY -8 SSLN-GVDKKLLSA-EVEKMLVQKG-----PNEGIEVVFGLLLYALA 58  
Db 29 SSVNLHIDPTLITMNDVROLGTQQNALNRDLFLTLKGKHPNLPKDKDFH-AMWLYR 88  
QY 59 ARTTS-----PKVORAD---SDVIF-SNSFGERNVVVTEGLKKVLDGC 96  
Db 89 VKSSSLQSDDDATGITYTREGVEVDLPKLTWDVFNKSGIGNR----- 132  
QY 99 APLTRFTNKLRTFGRTTEAYVDFCIAYKHKLPQLNAAAEGLGIPAEISYLAADFGLGCPK 158  
Db 133 -----TNALRVGRTDIALYLAFC-RQNRNLSYGGRLPDAGIPAGVHYLCADFL-TGAG 184  
QY 159 LSELQSRKQFASMYALKTEGGVWNTPVSNLRQLGR 194  
Db 185 LTDLECAVYIOAKQLK-KRGADDDVVTVNVRQLGK 219

Search completed: November 7, 2003, 14:49:49  
Job time : 36 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - nucleic search, using frame\_plus\_p2n model

run on: November 7, 2003, 20:42:48 ; Search time 2044 Seconds  
(without alignments)  
2354.347 Million cell updates/sec

title: US-09-613-486-15  
sequence: 1 MEIUMSDNLSNLVITASSL.....GGVNTPTVSNLRQLGRREV 198

scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

searched: 22781392 seqs, 1252238056 residues

total number of hits satisfying chosen parameters: 45562784

minimum DB seq length: 0  
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

>command line parameters:  
>DB=frame\_plus\_p2n.model -DBV=xlh  
>=/cgr2:/usrpro.spool/US09613486/runat\_07112003\_120411\_27105/app\_query.fasta\_1.391  
>B=EST -QFMT=fastap -SUFFIX=st -MINMATCH=0.1 -LOOPCL=C -LOOPEXT=0  
>INITs=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human43.cdi -LIST=45  
>OCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
>U=FMT=ptc -NCRW=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
>SPR=US09613486 @CGN\_1.1 2810 @runat\_07112003\_120411\_27105 -NCPU=6 -ICPU=3  
>O=MMAP -LARGEQUERY -NEG\_SCORES=C -WAIT -DSPLOCK=100 -LONGLOG  
>EV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
>GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DEEXT=7

tabase : EST:

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estm:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estto:\*  
8: em\_hrc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hrc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fur:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rnd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	88.5	8.9	706	13	BQ986943	BQ986943 QGF10N21.
2	88.5	8.3	748	28	BZ051926	BZ051926 jn57hr07.
3	87	8.8	477	12	BM132093	BM132093 TGSETRYB0
4	85.5	8.6	562	13	BQ459495	BQ459495 HA08017r
5	85.5	9.6	666	13	BUC07723	BUC07723 QGHSE15.Y
6	85.5	8.6	671	12	BJ433495	BJ433495 BJ433495
7	85.5	8.6	704	29	B2434184	B2434184 BONPR91TF
8	85.5	8.6	713	9	AU034876	AU034876 AU034876
9	85.5	8.6	751	12	BJ376355	BJ376355 BJ376355
10	85	8.6	654	13	BQ918284	BQ918284 QHB17604.
11	84.5	8.5	610	13	BQ848646	BQ848646 QGA7M01.Y
12	84.5	8.5	611	13	BQ984758	BQ984758 QGE5A05.Y
13	84.5	8.5	632	10	BG526441	BG526441 61-95 Ste
14	84.5	8.5	648	10	BG523153	BG523153 29-48 Ste
15	84.5	8.5	695	13	BQ993014	BQ993014 QGF27R14.
16	84	8.5	589	13	BQ982386	BQ982386 QGB16E23.
17	84	8.5	600	10	B5598403	B5598403 sv17b6.y
18	84	8.5	732	28	BH976504	BH976504 cdh67e01.
19	84	8.5	1098	10	BG247808	BG247808 602359515
20	83.5	8.4	633	28	BH096256	BH096256 RPCI-24-2
21	83.5	8.4	644	12	BJ340622	BJ340622 BJ340622
22	83.5	8.4	644	12	BJ435395	BJ435395 BJ435395
23	83.5	8.4	704	14	C921323	C921323 C921323.DIGT
24	83	8.4	600	14	C8343251	C8343251 CA32EN630
25	83	8.4	511	10	BQ421690	BQ421690 60249789
26	82.5	8.3	661	28	AQ949879	AQ949879 Sheared D
27	82.5	8.3	762	14	CB569015	CB569015 AGENCOURT
28	82	8.3	607	12	B1773596	B1773596 rs31h33.Y
29	82	8.3	627	13	BQ026578	BQ026578 QHG17E05.
30	82	8.3	750	13	BQ491813	BQ491813 PV_GEA012
31	82	8.3	528	13	BU410316	BU410316 603158539
32	82	8.3	523	14	CA582924	CA582924 ES-002599
33	82	8.3	1028	13	BU202935	BU202935 603102659
34	81.5	8.2	428	10	BG544734	BG544734 E2700 Chi
35	81.5	8.2	560	23	CG155340	CG155340 CSJ-K34.1
36	81.5	8.2	606	28	AQ09289	AQ09289 RPCI-23-4
37	81	8.2	523	28	AQ951071	AQ951071 Sheared D
38	81	8.2	727	28	BH929686	BH929686 cdh98c12.
39	81	8.2	807	29	B2577160	B2577160 rsn2_5276
40	80.5	8.1	836	12	BI085103	BI085103 602870643
41	80	8.1	432	13	BQ765335	BQ765335 sas16f01.
42	80	8.1	717	13	BQ969618	BQ969618 QHB39B11.
43	80	8.1	751	14	CA316784	CA316784 UI-M-FWC-
44	80	8.1	868	28	AZ691220	AZ691220 ENTJ241TF
45	80	8.1	999	12	BI199349	BI199349 602759615

# ALIGNMENTS

RESULT 1  
BQ986943  
LOCUS  
DEFINITION BQ986943 706 bp mRNA linear EST 21-AUG-2002  
QGF10N21.Yg.ab1 QG EFGHJ lettuce serriola Lactuca sativa cDNA clone  
QGF10N21, mRNA sequence.  
ACCESSION BQ986943  
VERSION BQ986943.1 GI:22404468  
KEYWORDS EST.  
SOURCE Lactuca sativa  
ORGANISM Lactuca sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;

Cichorieae; Lactuca,  
1 (bases 1 to 706)  
REFERENCE  
AUTHORS  
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,  
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, C., Ellison,  
P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,  
Church, S., Jackson, L., and Bradford, K.  
Lettuce and Sunflower ESTs from the Composite Genome Project  
http://comgenomics.ucdavis.edu/  
TITLE  
JOURNAL  
COMMENT  
Contact: Alexander Kozik (R.W.Michelmore)  
Department of Vegetable Crops, R.W.Michelmore Lab  
University of California at Davis (UCD)  
Amsundson Hall, UCD, Davis, CA 95616, USA  
Tel: 1-(530)-742-1742  
Fax: 1-(530)-752-9659  
Email: akozik@ucdavis.edu (michelmore@vegmail.ucdavis.edu)  
belongs to contig QG\_CA\_Contig108, see http://cgdb.ucdavis.edu/  
for details.  
Plate: QGF10 row: N column: 21.

FEATURES  
source  
1..706  
/organism="Lactuca sativa"  
/mol\_type="mRNA"  
/cui\_var="L.serriola"  
/db\_xref="taxon:4236"  
/clone="QGFI0N21"  
/lab\_host="E.coli"  
/clone\_lib="QG\_EFGH3 lettuce serriola"  
/note="Vector: pBRCDNASFIAB; The library was constructed  
from 10 different sources of RNA from a single genotype.  
Separate cDNAs were generated using primers that  
incorporated unique 5' and 3' tags to distinguish each  
source of RNA. cDNAs were then pooled, size-fractionated,  
directionally cloned into a custom medium-copy vector and  
transformations made with four size classes to minimize  
size bias. Details of each source of RNA and library  
construction can be obtained at http://cgdb.ucdavis.edu/  
TAG LIB=QG\_EFGH3 lettuce serriola  
TAG\_TISSUE=germinating seeds  
TAG\_SEQ=TCTGTGGG3"  
SE COUNT 195 a 110 c 190 g 211 t  
IGIN  
Alignment Scores:  
Seq. No.: 0.882 Length: 706  
ore: 88.50 Matches: 51  
Percent Similarity: 41.36% Conservative: 28  
Best Local Similarity: 26.70% Mismatches: 53  
Query Match: 8.93% Indels: 59  
Gaps: 13  
-09-613-486-15 (1-198) x BQ986943 (1-706)  
13 valleThAspAlaSerLeuAsnGlyValAspLysLeuSerAiaGluVal 32  
137 GTTTTACCAGAT---GGGTCCATACCATGATGATCCAGTCTCTGCGTCCAG---19C  
33 GluLysMetLeuValGlnCysGlyAlaProAsnGluGlyIleGluValValPheCysLeu 52  
191 ---CTGCACACACGAGGGGCTTCGGGAAGGTGT-----223  
53 LeuLeuTyAlaLeuAlaLeuAlaThrThrSerProLys-----ValGlnArg 68  
224 ---AGGAAGCTGCTCCAAAGATCTTGAGCCGATATGAGA 262  
69 AlaAsp-----SerAspValIlePheSer-----76  
263 GTTGAAGTTGTGACGCTGAGAACATTTGGGGGATGTGATGGGATTTGAATTCGAGA 322  
77 ---AsnSerPheGlyGluArgAsnValValValIleThrGluGlyAspLeuLys 92  
323 AGAGGGCAGATTAAATACCTTGTGTGATAAG-----CCTGGTGCATTT---364

QY 93 LysValLeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg 109  
Db 365 AAGGTGGTGGATTCCTTGGTCCACTTCGGGAAATGTTTACGTATGTGAGTACTTTAAGG 424  
QY 110 -----ThrPheGlyArg---ThrPheThrGluAlaTyValAspPheCysIleAlaTy 126  
Db 425 GGATTCACAAAGAGGAGAGCTCTTACACATGCAATAGCTAAATTTGATGTTGTGCT 484  
QY 127 LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAspSer 146  
Db 485 CAACATATTACAGAACAGCTTCTGCTCTAAAGAAAGAGCTGTGGCAGCTTGATCTTCA 544  
QY 147 -----TysLeuAla-----AlaAspPheLeuGlyThrCysPro 157  
Db 545 TCTTCATCTTGTGCTATTTAATGATAGGGGAATTTCTGATTTTCATCTTCATATGTCAA 604  
QY 158 LysLeuSerGlnLeuGlnSerArgLysMet 168  
Db 605 TATAGGAAAAAACAACAAACAAAGGATT 637

RESULT 2  
BZ051926/c  
LOCUS  
DEFINITION  
748 bp DNA linear GSS 09-OCT-2002  
Jcr57h07.g1 B.oleracea001 Brassica oleracea genomic, genomic survey  
sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Richard K. Wilson  
Genome Sequencing Center,  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Plate: jnr57 row: h column: 07  
Seq primer: -28RFPOT reverse  
Class: shotgun  
High quality sequence start: 56  
High quality sequence stop: 551.  
Location/Qualifiers  
1..748  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3712"  
/clone\_lib="B.oleracea001"  
/note="Vector: pOTw13; Whole genome shotgun library from  
preparing buds. DNA was purified from a crude nuclear  
prep using Brassica oleracea T01000DH3 buds provided by  
Thomas Osborn at the University of Wisconsin.  
DNA was provided by Pablo Rabiowicz (CSH) and the  
shotgun library prepared at Washington University Genome  
Sequencing Center."  
BASE COUNT 209 a 164 c 136 g 227 t  
CRIGIN

Alignment Scores:  
Seq. No.: 0.96 Length: 748  
ore: 88.50 Matches: 33  
Percent Similarity: 47.00% Conservative: 14  
Best Local Similarity: 33.00% Mismatches: 40  
Query Match: 8.93% Indels: 13  
Gaps: 5  
US-09-613-486-15 (1-198) x BZ051926 (1-748)



/dev stage="0-7 DAP (days after pollination)";

/lab\_host="x210-cold"

/clone\_lib="RA"

/notes=Vector: pluescript SK+; Site\_1: EcoRI (5'-end of cDNA); Site\_2: XhoI (3'-end of cDNA); 0-7 DAP(days after pollination) Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, Sall). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable. Average insert size is 1 kb

SE COUNT 145 a 83 c 181 g 153 c

IGIN

ignment Scores:

ed. No.: 1.54 Length: 562  
ore: 85.50 Matches: 40  
Percent Similarity: 40.00% Conservativeness: 38  
st Local Similarity: 20.51% Mismatches: 84  
ery Match: 8.63% Indels: 33  
Gaps: 8

-09-613-486-15 (1-198) x BQ459495 (1-562)

```

22 GlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAla 41
|||||
4 GGCATGGAGAGAGCGCATGGATTGCTACAGGAGTGCT----- 45
42 ProAsnGluGlyIleGluValValPheGlyLeuLeuTyrAlaLeuAlaAlaArgThr 61
|||||
46 ---GGAGGGGATCCAAAGGTGAGATTGGGGCTGTAGCTACAACTGGTGTGATGCG 102
62 ThrSerProLysValGlnA-GAlaAspSerAspValIlePheSerAsnSerPheGlyGlu 91
|||||
103 CTTCGGAGGAGCGGAGGTAGATGATGCGATCGAGTTGTTGATAGATGTGCAAGAG 162
82 -----AfgAsnValValIleThrGluGlyAspLeuLysLysValLeuAspGly 97
|||||
163 CATGACCTCCACAGGAGATTGCTGTAATCTTGAGTCTTAAATGATGATGATGCT 222
98 CysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGlu 117
|||||
223 TACTGCGGTGTGGAAAGGTTCCAGATGCAATTGAGGTGTTTGGCAAAATGGGTGAGAAG 282
118 AlaTyrValAspPheCysIleAlaTyr-----LysHisLys 129
|||||
283 AGCTGCACACAGATGCACTCTCAATAATAGTCTGATTACTGCTGGGGAAGACAAA 342
130 LeuProGlnLeuAsnAlaAlaAlaGluLeu-----GlyIle--- 141
|||||
343 CTT-----GTGTGAGGCGAAGAATTGTACAAGGAGATGGGGAGCGTGGTACTTAAT 396
142 ProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGlu 161
|||||
397 CCGACGAATACACTTATGCTTA-----CTCATGAGTCTGCTTCAGCGTTGATAAT 450
162 LeuGlnGlnSerArgLysMetPheAlaSerVetTyr-----AlaLeuLysThrGluGly 179
|||||
451 GTGATGACGCTGTGTTGTTACTTCAGCAAGATGTTGATGTTGCTTCAGGCCCAATGCC 510
180 GlyValValAsnThrProValSerAsnLeuArgGlnLeuGlyArg 194
|||||
511 AATGCTTTAAACAAAGTCATAAGTGGCTTGGTGAAGGTGATAGG 555

```

SULT 5

007723

CUS

FINITION

QGH5E15.Yg.ab1 QG\_EFGHJ lettuce serricola Lactuca sativa cDNA clone

QGH5E15, mRNA sequence.

BU007723

BU007723.1

GI:22442118

EST.

SOURCE  
ORGANISM  
Lactuca sativa  
Lactuca sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;  
Cichorieae; Lactuca.  
REFERENCE  
1 (bases 1 to 666)  
AUTHORS  
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, X., Rieseberg, L.,  
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison,  
P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z.,  
Church, S., Jackson, L., and Bradford, K.  
TITLE  
Lettuce and Sunflower ESTs from the Compositae Genome Project  
http://compgenomics.ucdavis.edu/  
JOURNAL  
UNPUBLISHED  
CONTACT: Alexander Kozik [R.W.Michelmore]  
Department of Vegetable Crops, R.W.Michelmore Lab  
University of California at Davis (UCD)  
Armstrong Hall, UCD, Davis, CA 95616, USA  
Tel: 1-(530)-742-1742  
Fax: 1-(530)-752-9659  
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]  
belongs to contig QG\_CA\_Contig108, see http://cgpdb.ucdavis.edu/  
for details.  
Plate: QGH5 row: E column: 15.  
FEATURES  
Location/Qualifiers  
1..666  
/organism="Lactuca sativa"  
/mol\_type="mRNA"  
/cultiivar="L.serricola"  
/db\_xref="taxon:4236"  
/clone="QGH5E15"  
/lab\_host="E.coli"  
/clove\_lib="QG\_EFGHJ lettuce serricola"  
/note="Vector: pBRCDNASFIAB; The library was constructed  
from 10 different sources of RNA from a single genotype.  
Separate cDNAs were generated using primers that  
incorporated unique 5' and 3' tags to distinguish each  
source of RNA. cDNAs were then pooled, size-fractionated,  
directionally cloned into a custom medium-copy vector and  
transformations made with four size classes to minimize  
size bias. Details of each source of RNA and library  
construction can be obtained at http://cgpdb.ucdavis.edu/  
TAG LIB=QG\_EFGHJ lettuce serricola  
TAG TISSUE=chemical induction  
TAG\_SEQ=TCGTAGCCGG"

BASE COUNT 176 a 115 c 191 g 183 t 1 others  
ORIGIN

Alignment Scores:

Pred. No.: 1.98 Length: 666  
Score: 85.50 Matches: 50  
Percent Similarity: 40.78% Conservativeness: 23  
Best Local Similarity: 27.93% Mismatches: 47  
Query Match: 8.63% Indels: 59  
Gaps: 13

US-09-613-486-15 (1-198) x BU007723 (1-666)

QY 13 ValIleThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuSerAlaGluVal 32  
Db 184 GTTTTACCGCAT---GGGTCTCTACCATGATGTGATCCAGTGTCTTGGGTCCAG--- 237  
QY 33 GluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeu 52  
Db 238 -----CTGGCAGCAGGCGGCGTTTCGGGAAGGTGTT----- 270  
QY 53 LeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLys-----ValGlnArg 68  
Db 271 -----AGGAAGCTGCTCCAAAGATCTTGGAGCCGATTTATGAGA 309  
QY 69 AlaAsp-----SerAspValIlePheSer----- 76  
Db 310 GTTGAAGTTGTGACGCTGAAGACATTTGGGGATGTGATTTGGGGATTTGAATTCGAGA 369



```

77 -----AsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLys 92
370 AGAGGGCAGATTAATAGCTTGGTGATAAG-----CCTGGGACTT--- 411
93 LysValLeuAspGlyCysAlaProLeuThrArgPhe-----AtgPheThrAsnLysLeuArg 109
412 AAGGTGGTGATCTTTGGTCCCACTTCGGGAATGTTTCAGTATGCGAGTACTTTAAGG 471
110 -----ThrPheGlyArg-----ThrPheThrGluAlaTyrValAspPheCysIleAlaTyr 126
472 GGAATGACAAAAGGAGAGAGCTTCTTACCAATGCAATTAGCTAAATTTGATGTTGGCCT 531
127 LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAspSer 145
532 CAACATATTCAAGAACGAGCTTCTCTGCTGTANAGAGAAGCTGTGGCAGCTTGATCTTCA 591
147 -----TyrLeuAla-----AlaAspPheLeuGlyThrCys 156
592 TCTTCATCTTTGTGCTATTATATGATAGGGGAATTTCTGATTTTCATCTTCATATGT 648

RESULT 6
BJ433495/c
LOCUS
DEFINITION
BJ433495 Dictyostelium discoideum cDNA library, VF Dictyostelium
discoideum cDNA clone ddv22f05 3', mRNA sequence.
BJ433495.1 GI:19408217
EST.
DICTYOSTELIUM DISCOIDEUM
SOURCE
ORGANISM
Eukaryota; Eukaryota; Dictyostelium; Dictyostelium.
Eukaryota; Mycetozoa; Dictyostelium; Dictyostelium.
1 (bases 1 to 671)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoideum at the vegetative
stage
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1..671
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv22f05"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"
19 a 129 c 96 g 227 t

Alignment Scores:
ed. No.: 2 Length: 671
ore: 85.50 Matches: 46
cent Similarity: 35.4% Conservative: 21
st Local Similarity: 24.34% Mismatches: 85
ery Match: 8.63% Indels: 37
3: 12 Gaps: 5

-09-613-486-15 (1-198) x BJ433495 (1-671)
33 GluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeu 52
663 GAGAGGTTAAGTTAGTGGTAAGACACCAATGAGAGATTGAGAGATTGGGAGTATTA 604
53 LeuLeuTyrAlaLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 72

```

```

Db 603 TCATCATCTTGTTCAGCTCATATGACA-----CAATTAACAAGTGAGAT 556
Qy 73 ValLLePheSerAsnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLys 92
Db 555 TTAGATTTCACACGACAAAGAGTGGTATTAATGTTGTACATTGTCGAGATCAAAATTTAAAG 496
Qy 93 LysValLeuAspGlyCysAlaProLeuThrArgPhe----- 104
Db 495 TTGGGTGTTTAAAGGTATTTCACGAGTTTCATAAGCTTTTAAAAACAAGGTAAACGTTTCA 436
Qy 105 -----ThrAsnLysLeuArgThrPheGlyArgThrPheThr 116
Db 435 GTTGGTACTGATCTCCCGCTAGTATGATGATTAGATATGTTGGTGAATTTACGCTACT 376
Qy 117 GluAlaTyrValAspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAla 136
Db 375 GCAGCTTACATTGATAAACTCTCTGCAAACTACTCAATTCATTT 334
Qy 137 AlaGluLeuGlyIleProAlaGluAspSerTyrLeu----- 148
Db 333 ---GAAGGTGGTGAACCTGTAAACACCATCATATAGATTTTGTCAATGGCACTATTAAC 277
Qy 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgCysMet 168
Db 276 GGTGCAAGGCAATTGGGTATCGATGATAAAGTTGGTTCACCTTCAAAATTCGTAAA----- 223
Qy 169 PheAlaSerMetTyrAlaLeuLysThrGluGlyValValAlaAsnThrProValSerAsn 188
Db 222 TTTGCAGATTTCATCGCTGTCAAAGTTTCAAGTCATCCAGTCTATGATCCAAATCTCTCAT 163
Qy 189 LeuArgGlnLeuGlyArgArgGluVal 197
Db 162 TTAGTTTATGTTGGTACTAATCATGTC 136

RESULT 7
BZ434184/c
LOCUS
DEFINITION
BONFR91TF BO_1.6_2_KB tot Brassica oleracea genomic clone BONFR91,
genomic survey sequence.
BZ434184
BZ434184.1 GI:26684219
SOURCE
GSS.
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 704)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other_GSSs: BONFR91TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..704
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BONFR91"
/clone_lib="BO_1.6_2_KB tot"
/note="Vector: pHOSt1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOSt1 using BstXI linkers"
169 a 168 c 144 g 223 t

BASE COUNT
ORIGIN

```

```

alignment Scores:
  ed. No.:      2.15      Length:      704
  Ore:          85.50      Matches:      49
  Percent Similarity: 38.16%      Conservative: 30
  Best Local Similarity: 23.67%      Mismatches: 81
  Query Match:      8.63%      Indels:      47
  DB:              29      Gaps:      8

-09-613-486-15 (1-198) x BZ434184 (1-704)
7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLys 26
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
574 TCTGACCTCAACACAGTACATGCACTTCCCTCTCATCTCAAGGCTTTGGAGAGTCAA 515
27 Leu-----LeuSerAaGluValGlu 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
514 TTTGCTTCTATGCTTCTATCTCCAGCAGCAATGGGGTCTCACCAGGGAATCAGAA 455
34 LysMetLeuValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeuLeu 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 AAG-----AACCCCAAGGAGTCTTGCAATGTTGTTCTCTCC----- 419
54 LeuTyrAlaLeuAlaAaArgThrThrSerProLysValGlnArgAlaAspSerAsp--- 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 -----ACTACTCTCCAGAGATTGAGT-GAGTGATCATGACAAA 380
73 -----ValIlePheSerAsnSerPheGlyGlu---ArgAsnVal 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
379 GAGGAGGATGAGATTGAAGACTCGTATTGGAACTAGTTTGGGGAAGTTGAGAGATT 320
85 ValValThrGluG:YAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPhe 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
319 GTTGTGGCCACAGCTGAAGACACAGATTGTGATGCAC---GCTGCAAGGAAGTTGAGCA 263
105 ThrAsnLysLeuAaGThrPheGlyArgThrPheThrGluAlaTyrValAspPheCysIle 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
262 ACAANTCTGCAAGAGAGCTGACACAAA-----GCTGAGAACAAGTTGAGAGAGAGCT 209
125 AlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGlu 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
208 GACAAACAGCTGAAAAGAGATAAGTAAGAGAGCCACTGAGGTGAGCTATACCCCTAT 149
145 AspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGln 164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 GATAAGCTCCCTTTCCCCCAAGAGTTCTCCACCAAGCTCAGACGAAGGTGCTCTCCAAG 89
165 SerArgLysMetPheAlaSerXerTyrAlaLeuLysThrGluGlyGlyValValAsnThr 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
88 TTCAGGAAGAGATCTTAGTGATGTT-----GGGTGAGGCTTCCA 50
185 ProValSerAsnLeuArgGln 191
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
49 GAAATCTCGAGTATCGGTGAA 29

A0034876 713 bp mRNA linear EST 28-APR-1999
A0034876 Dictyostelium discoideum SL (H. Urushihara) Dictyostelium
discoideum cDNA clone SL5629, mRNA sequence.
A0034876
CESSION
YWORDS
RSION
ORGANISM
Dictyostelium discoideum
Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
1 (bases 1 to 713)
MoriO.T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development

JOURNAL      DNA Res. 5 (6), 335-340 (1998)
MEDLINE      99156227
PUBMED      10048482
COMMENT      Contact: Hidexu Urushihara
              Institute of Biological Sciences
              University of Tsukuba
              1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
              Tel: 81-298-53-4664
              Fax: 81-298-53-6614
              Email: hidexu@biol.tsukuba.ac.jp
              PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
              POLYA=No.

FEATURES
    source
        1..713
            /organism="Dictyostelium discoideum"
            /mol_type="mRNA"
            /strain="AX4"
            /db_xref="taxon:44689"
            /clone="SL5629"
            /dev_stage="slug"
            /clone_lib="Dictyostelium discoideum SL (H.Urushihara)"

BASE COUNT  241 a 105 c 132 g 235 t
ORIGIN
Alignment Scores:
  Pred. No.:      2.19      Length:      713
  Score:          85.50      Matches:      45
  Percent Similarity: 35.43%      Conservative: 21
  Best Local Similarity: 24.34%      Mismatches: 85
  Query Match:      8.63%      Indels:      37
  DB:              9      Gaps:      5

US-09-613-486-15 (1-198) x A0034876 (1-713)
QY 33 GluLysMetLeuValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeu 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 40 GAAGAGGTTAAGCTTAGTGGTAAGAGACCAATTGAGAGATTGAGAGATTGGGAGTATT 99
QY 53 LeuLeuTyrAlaLeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAsp 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 100 TCATCATCTATTGATGCTCATATGACA-----CAATTAAACAGTGAAGAT 147
QY 73 ValIlePheSerAsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLys 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 148 TTAGATTTCACAGCAAGAGAGTGGTATTAAATGTTGATACATTGCCAGATCAATTTAAG 207
QY 93 LysValLeuAspGlyCysAlaProLeuThrArgPhe----- 104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 208 TTGGGTGTTTAAAGGTATTTCACCAAGTTCATAAACTTTTAAACAAGGTGTAAACGTTTCA 267
QY 105 -----ThrAsnLysLeuAaGThrPheGlyArgThrPheThr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 268 GTTGTACTGATCTGCGCGCTAGTAAATGATGATTTAGATATGTTGGTGAATACGTACT 327
QY 117 GluAlaTyrValAspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAla 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 328 GCAGCTTACATGATAAACTCTGCAAAATACATCATCAATT----- 369
QY 137 AlaGluLeuGlyIleProAlaGluAspSerTyrLeu----- 148
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 370 ---GAAGGTGGTGAACCTGTAAACCATCATATAAGATTTTGTCAATGGCAACTATTAAAC 426
QY 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMet 168
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 427 GGTGCAAGGCATTTGGGTATCGATGATAAGTGGTTCCTCAATTTGGTAAA----- 480
QY 169 PheAlaSerXerTyrAlaLeuLysThrGluGlyGlyValValAsnThrProValSerAsn 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 481 TTTGCAAGATTTCATCGCTGTCAAAGCTTTCAAGTCATCCAGTCTATGATCAATCTCTCAT 540
QY 189 LeuArgGlnLeuGlyArgArgGluVal 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 541 TTAGTTTATGTTGGTACTAATCATGTC 567
```

```

SULT 9
376355/C
CUS
FINITION
ACCESSION
BJ376355
KEYWORDS
EST
SOURCE
Dictyostelium discoideum
ORGANISM
Dictyostelium discoideum
REFERENCE
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
AUTHORS
Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i.T.
TITLE
Full length cDNA of Dictyostelium discoideum at the culmination
stage
JOURNAL
Unpublished
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tsun@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
1..751
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddc28n02"
/dev_stage="Culmination stage"
/clone_lib="Dictyostelium discoideum cDNA library, CF"
BASE COUNT 244 a 139 c 104 g 264 t
IGIN
Alignment Scores:
ore: 2.36 Length: 751
ore: 85.50 Matches: 46
cent Similarity: 35.45% Conservative: 21
st Local Similarity: 24.34% Mismatches: 95
ery Match: 6.63% Indels: 37
Gaps: 5
-09-613-486-15 (1-198) x BJ376355 (1-751)
33 GLuLysXetLeuValGlnLysGlyAlaProAsnGluGlyIleGluValPheGlyLeu 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
663 GAAGAGGTAAAGGTTAGGTGGTGAAGAGACCAATTGAGAGATTTCGAGATTA 604
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 LeuLeuTyrAlaLeuAlaAlaArgThrSerProLysValGlnArgAlaAspSerAsp 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
603 TCATCATCATTCATTGCAGCTCATATGACA-----CAATTAAACAGTGAAGAT 556
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 ValIlePheSerAspSerPheGlyGluArgAsnValValThrGluGlyAspLeuLys 92
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
555 TTAGATTTTCACAGCAAGAGTGGTATTAAATGTTGACATTGTCAGAGTCAAAATTAAAG 496
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 LysValLeuAspGlyCysAlaProLeuThrArgPhe----- 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
495 TTGGGTGTTAAAGGTATTCACCACTTTATAACTTTTAAACACAGGTGTAACGTTTCA 436
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
105 -----ThrAsnLysLeuArgThrPheGlyArgThrPheThr 126
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
435 GTTGCTACTGATTTCGCGCTAGTAAATGATGATTAGATGTTGGGTGAATTACGTA 376
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 GluAlaTyrValAspPheCysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAla 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
375 GCAGCTTACATTGATAAACTCTCTGCAATACTCATTCATT----- 334
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
137 AlaGluLeuGlyIleProAlaGluAspSerTyrLeu----- 148
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

333 ---GAAGGTGGTGAACCTGTAAACACCATCATATAAGATTTTGTCAATGGCAACTATTAAC 277
149 A..AlaAspPheLeuGlyThrCysProLysLeuSerG..LeuGlnSerArgLysMet 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
276 GTGCGCAAGCATGTGGTATCGCATGATAAAGTTGTTCAATTCGTTAA----- 223
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 PheAlaSerMetTyrAlaLeuLysThrGluGlyGlyValValAsnThrProValSerAsn 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 TTGCGAGATTTCATCGCTGTCAAAGTTTCAAGTCATCCAGTCTATGATCCATCTCTCAT 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
189 LeuArgGlnLeuGlyArgArgGluVal 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
162 TTAGTTTATGTTGGTACTAATCATGTC 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 10
BO916284
LOCUS
DEFINITION
BO916284
ACCESSION
BO916284
VERSION
BO916284.1 GI:22315065
KEYWORDS
EST
SOURCE
Helianthus annuus (common sunflower)
ORGANISM
Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroidae;
Heliantheae; Helianthus.
REFERENCE
1 (bases 1 to 654)
AUTHORS
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevaller,P., Ziegler,J., Ellison
,P., Kolkman,S., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
Church,S., Jackson,L. and Bradford,K.
TITLE
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://comgenomics.ucdavis.edu/
JOURNAL
Unpublished
COMMENT
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Rasmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]
belongs to contig OH_CA_Contig3046, see http://cgpdb.ucdavis.edu/
for details.
Plate: QHB17 row: G column: 04.
FEATURES
Location/Qualifiers
1..654
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QHB17G04"
/lab_host="E.coli"
/clone_lib="QH_ABCDI sunflower RHA801"
/note="vector: pBRCDNASFIAB; The library was constructed
from 11 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB=QH_ABCDI sunflower RHA801
TAG_TISSUE=shoots environmental stress
TAG_SEQ=TCGCACACGG"
BASE COUNT 201 a 110 c 175 g 168 t
ORIGIN
Alignment Scores:
Pred. No.: 2.24 Length: 654
Score: 85.00 Matches: 41

```







